LUND UNIVERSITY

Optical Multiplexing for Ultrafast Videography

Ek, Simon

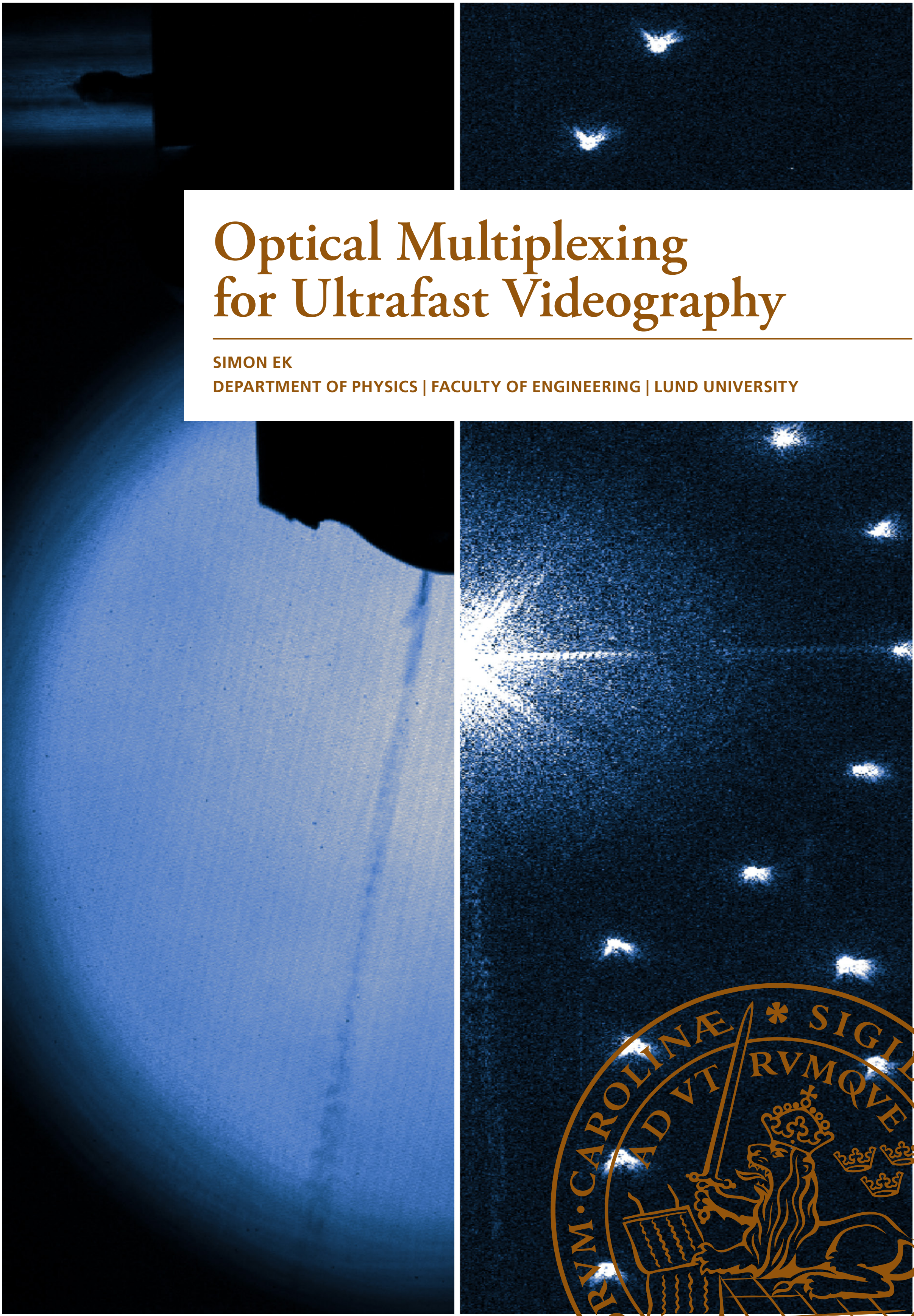
2025

Citation for published version (APA):
Ek, S. (2025). *Optical Multiplexing for Ultrafast Videography*. Department of Physics, Lund University.

Total number of authors:
1

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00



Optical Multiplexing for Ultrafast Videography

SIMON EK

DEPARTMENT OF PHYSICS | FACULTY OF ENGINEERING | LUND UNIVERSITY

Optical Multiplexing for Ultrafast Videography

Optical Multiplexing for Ultrafast Videography

by Simon Ek

LUND
UNIVERSITY

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

To be presented, with the permission of the Faculty of Engineering of Lund University, for public criticism in the Rydberg Lecture Hall (Rydbergssalen) at the Department of Physics on Friday, the 9th of May 2025 at 13:15.

DOKUMENTDATABLAD enl SIS 61 41 21

Organization
LUND UNIVERSITY

Department of Physics
Box 117
SE–221 00 LUND
Sweden

Document name
DOCTORAL DISSERTATION

Date of disputation
2025-05-09

Sponsoring organization

Author(s)
Simon Ek

Title and subtitle
Optical Multiplexing for Ultrafast Videography

Abstract

Capturing ultrafast phenomena requires imaging techniques capable of resolving events occurring on femtosecond time scales. Traditional high-speed cameras are limited by electronic readout speeds and cannot reach the terahertz (THz) frame rates necessary to observe processes such as laser filamentation, plasma formation, or light-matter interactions in general. To overcome these limitations, this thesis explores and advances Frequency Recognition Algorithm for Multiple Exposures (FRAME), an optical multiplexing technique that encodes multiple frames within a single exposure by using spatially modulated illumination pulses. These encoded frames are later separated computationally, enabling single-shot ultrafast videography.

A central achievement of this work is the first demonstration of phase-sensitive FRAME, which significantly enhances the technique's ability to detect small refractive index variations. This was then extended to phase-sensitive ultrafast FRAME, allowing the visualization of laser filamentation in a single shot with high levels of detail. By leveraging FRAME's ability to extract phase information directly from the recorded data, this work shows that it is possible to reconstruct highly detailed image sequences of plasma formation and filament propagation with FRAME.

This thesis also investigates the fundamental limitations and capabilities of FRAME, including its maximum achievable sequence depth, sensitivity, and resolution trade-offs. FRAME's flexibility is highlighted through applications in diverse experimental configurations, from fluorescence-based multispectral imaging to high-speed shadowgraphy of fluid dynamics.

The results place FRAME within the broader historical context of imaging technology, illustrating how ultrafast optical multiplexing represents a paradigm shift in high-speed imaging. While challenges such as image quality and sensitivity remain, FRAME provides a versatile, easy-to-implement approach that pushes the limits of ultrafast videography. Future work should focus on refining the resulting image quality and improving ease of use, ultimately enabling FRAME's adoption in broader scientific fields.

Key words
Ultrafast Videograpy, Optical Multiplexing, Photonics, Structured Light, FRAME, Optical Diagnostics

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language
English

ISSN and key title
ISSN: 1102-8718

ISBN
978-91-8104-481-2 (print)
978-91-8104-482-9 (pdf)

Recipient's notes

Number of pages
224

Price

Security classification

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources the permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature ____________________ Date 2025-04-15

Optical Multiplexing for Ultrafast Videography

by Simon Ek

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: A detector image of a fuel injection event and its corresponding Fourier transform, illustrating the same information in two different representations.

Cover illustration back: A stitched together image of a laser induced plasma event. Slava Ukraini!

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: 978-91-8104-481-2 (print)
ISBN: 978-91-8104-482-9 (pdf)
ISSN: 1102-8718

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

NORDIC SWAN ECOLABEL
Printed matter
3041 0903

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

Dedicated to my children
Ebbe – Elsa – Signe

Table of Contents

Abstract

Capturing ultrafast phenomena requires imaging techniques capable of resolving events occurring on femtosecond time scales. Traditional high-speed cameras are limited by electronic readout speeds and cannot reach the terahertz (THz) frame rates necessary to observe processes such as laser filamentation, plasma formation, or light-matter interactions in general. To overcome these limitations, this thesis explores and advances Frequency Recognition Algorithm for Multiple Exposures (FRAME), an optical multiplexing technique that encodes multiple frames within a single exposure by using spatially modulated illumination pulses. These encoded frames are later separated computationally, enabling single-shot ultrafast videography.

A central achievement of this work is the first demonstration of phase-sensitive FRAME, which significantly enhances the technique's ability to detect small refractive index variations. This was then extended to phase-sensitive ultrafast FRAME, allowing the visualization of laser filamentation in a single shot with high levels of detail. By leveraging FRAME's ability to extract phase information directly from the recorded data, this work shows that it is possible to reconstruct highly detailed image sequences of plasma formation and filament propagation with FRAME.

This thesis also investigates the fundamental limitations and capabilities of FRAME, including its maximum achievable sequence depth, sensitivity, and resolution trade-offs. FRAME's flexibility is highlighted through applications in diverse experimental configurations, from fluorescence-based multispectral imaging to high-speed shadowgraphy of fluid dynamics.

The results place FRAME within the broader historical context of imaging technology, illustrating how ultrafast optical multiplexing represents a paradigm shift in high-speed imaging. While challenges such as image quality and sensitivity remain, FRAME provides a versatile, easy-to-implement approach that pushes the limits of ultrafast videography. Future work should focus on refining the resulting image quality and improving ease of use, ultimately enabling FRAME's adoption in broader scientific fields.

List of publications

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

I **Long sequence single-exposure videography using spatially modulated illumination**
S. Ek, V. Kornienko, E. Kristensson
Scientific Reports, 10 (1) pp. 18920, 2020

II **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

III **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

IV **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

V **Ultrafast single-shot imaging of laser induced plasma events in air**

S. Ek, Y. Bao, V. Kornienko, E. Kristensson
Manuscript

VI **Coherence lifetime imaging**

V. Kornienko, Y. Bao, **S. Ek**, A. Hosseinnia, M. Reaveesh, C. J. Kliewer, J. Bood, A. Ehn, E. Kristensson
Manuscript

Acknowledgements

When I was three years old, I allegedly gave my first oral presentation. According to the story, all of us children at Förskolan Blomrankan were asked to find an object at the beach in Lomma and prepare a short presentation about it for the following day. I chose a seashell and likely told the other children that it had once been the home of a mollusc, which was now in mollusc heaven. I do not remember this episode myself, but I have been told the story on several occasions. In retrospect, I find it illustrative of the encouragement and trust we were shown by our teachers from a very early age. For this, I am sincerely grateful. If I have acquired any proficiency in the art of rhetoric, it began at Blomrankan, under the caring supervision of Susanne, Åsa, Charlotte, and Karin.

My current, and likely final, supervisor Elias Kristensson has now brought to completion what was begun more than 30 years ago—by guiding me through the final stage of the Swedish education system. You hired me in 2019, which in itself was a significant act of trust, and since then you have continuously shown confidence in my ability to work independently, while still providing guidance when needed. I have even been encouraged to pursue directions not directly tied to your original research plan. Ultimately, it is you, not I, who must (figuratively) stand before the great judge — the research funding organizations — and account for how the realized research aligns with the funded proposal. Given that the outcome of that evaluation, and to some degree your own future career, depends on the work I have carried out during these years, I am deeply grateful for the trust you have shown me. I also want to thank you for your honest feedback and for always being available. Although our scheduled meetings have been relatively infrequent, I have always had the impression that you've been aware—more or less—of what I've been up to, and I've found it easy to reach out for further advice when needed. Our regular, unscheduled encounters in the dressing rooms at Gerdahallen have also served as excellent opportunities to discuss the progress of my work.

The colleagues I have worked closest with are, without a doubt, Dr Vassily Kornienko and Dr Yupan Bao. Vassily and I were both offered our PhD positions around the same time, while still working on our Master's theses at the Division of Atomic Physics. I still remember the joy I felt upon realizing that I would get to continue working alongside Vassily for at least another five years. I have asked you for advice a countless number of times and your positive attitude have really helped me stay motivated in my own work. I particularly what to thank you for proofreading and commenting on every chapter of this thesis. Yupan, it has been an absolute pleasure working with you in the lab. I especially appreciated our non-research-related conversations, during which we often shared a similar view on life and the world events that have been unfolding around us. I also want to say a special thanks to David Andersson for your contributions in our group discussions and for your company in Vienna.

As customary, I want to thank my parents, Eva and Alf Ek, for their unwavering support and general parenting. In particular, I am deeply grateful for the way you have supported me and my own family during these years. My mother deserves a special mention for taking care of my children every Tuesday afternoon for the past ten years. This has allowed me to work late at least once a week—something that has been especially valuable during the more intense periods of writing this thesis. I can never truly repay you for this, but the offer remains: you and dad can live in a tent in my garden when you get old. I will also do my best to be an equally present grandfather, should that day ever come. [1]

Hanna, my love, thank you for being there for me. To be completely honest I could probably have done this thesis without you, but life in general I could not. You are the hub around which our family revolves. We all agree that mum is the best. Forgive me for making you read this lengthy book.

Finally, thanks be to God for creating a world so complex that we will never fully understand it, yet so comprehensible that meaningful progress can be made—and for endowing us with the ability and curiosity to seek that understanding.

---

[1] Dad, please translate for mom.

Popular summary

In the human world, lead times are long, and everything seems slow. In the vastness of the universe, changes occur over millions of years, making time feel almost frozen. But if we turn our attention to the fascinating world of the microscopic, particularly to interactions between light and matter, time scales become incredibly short—both due to the unmatched speed of light and the small length scales involved. These processes are dynamic, complex, and sometimes even chaotic. Understanding them in detail is challenging but crucial for advancing technologies such as next-generation particle accelerators and laser-based lightning conductors.

One example of light-matter interaction is laser filamentation. When an intense laser pulse enters a gas, the gas's optical properties change instantaneously, slowing down the speed of light—what we call an increase in the refractive index. Since the pulse is typically most intense at its center, the light slows down the most there. The effect can be visualized with a simple analogy: imagine three friends walking hand in hand, casually blocking a bike lane. If the person in the middle suddenly slows down while the others maintain their pace, the entire group bends inward. Similarly, light bends toward the center, where intensity—and therefore the slowing effect—is strongest. This phenomenon is called self-focusing, and it causes the pulse to become even more intense. The process reinforces itself until the pulse becomes so intense that it starts tearing electrons away from the gas atoms. The free electrons cause the laser pulse to stop focusing and instead begin expanding. As intensity decreases, the system reverts to self-focusing, and the cycle repeats. The result is a dynamic balance where the pulse diameter fluctuates slightly but overall remains very small.

To film how this laser pulse moves with a spatial resolution matching its thickness (about 30 micrometers) requires an exposure time of 100 femtoseconds (0.000 000 000 000 100 seconds). The corresponding frame rate is 10 terahertz (THz), or 10 trillion frames per second. For comparison, a typical computer processor operates at 3 gigahertz (GHz)—three thousand times slower. A camera sensor, in many ways similar to a processor, has a maximum speed that makes filming at 1 THz or more completely impossible with conventional high-speed cameras. A different approach is needed.

The solution is ultrashort illumination pulses. Even if the camera's shutter is technically open for a longer period, light only reaches the sensor during the few femtoseconds that the illumination pulse lasts. These pulses, often around 30 femtoseconds long, resemble extra thin pancakes, with a thickness about a tenth of the width of a human hair (10 micrometer). Since the flash of light itself is so incredibly short, it effectively ”freezes” the event, even if the sensor itself is comparatively slow.

Imagine viewing from the camera's perspective. From the left, a laser filament rushes forward at the speed of light. At the same time, a pancake-like laser pulse—the illumination pulse—travels toward you, passing through the filament. At the moment they intersect, an image of the filament is imprinted onto the illumination pulse, which is then captured by the camera. To record multiple frames, a whole train of short light pulses is generated, passing through the filament in rapid succession, with each pulse capturing a new phase of the process.

This setup, however, presents a practical problem: all illumination pulses overlap on the camera sensor, making it seemingly impossible to separate them. The solution is to assign each pulse a unique stripe pattern. These patterns can vary in frequency (stripe thickness) or angle. In the recorded image, the overlapping stripe patterns remain visible, and using a computer algorithm, it is possible to ”lock-in” to individual patterns and reconstruct their respective images. This works similarly to how a radio locks-in to a specific station by tuning to its exact frequency, filtering out all other channels.

During my PhD research, I have used this method to film laser filaments with extremely high temporal and spatial resolution. Additionally, I have helped refine the technique to make it more sensitive to changes in refractive index, which has sometimes been essential for obtaining any image at all. Altogether, my work represents another step toward truly ultrafast videography, allowing us to capture nature’s fastest phenomena and deepen our understanding of light-matter interactions on record-breaking time scales.

Populärvetenskaplig sammanfattning

I människornas värld är ledtiderna långa och det mesta tycks trögt. I universums makrokosmos mäts förändringar i miljontals år så att tiden tycks stå stilla. Om vi istället vänder blickarna mot det förunderliga mikrokosmos, specifikt mot växelverkan mellan ljus och materia, är tidsskalorna istället extremt korta. Både på grund av ljusets oöverträffade hastighet och de mycket korta längdskalorna. De här processerna är dynamiska och komplexa - närmast kaotiska. Att förstå dessa fenomen i detalj är svårt, men kan ha avgörande betydelse för utvecklingen av allt från nästa generations partikelacceleratorer till laserbaserade åskledare.

Ett exempel på ljus-materia-växelverkan är laserfilamentering. När en mycket intensiv ljuspuls träffar en gas förändras gasens optiska egenskaper omedelbart och på ett sådant sätt att ljusets hastighet saktar ner. Vi säger att brytningsindex ökar. Eftersom pulsen (oftast) är mest intensiv i mitten blir nedsaktningen av ljuset som störst där. Effekten kan förklaras med en enkel bild. Föreställ dig tre vänner som går hand i hand och blockerar en cykelbana. Plötsligt saktar personen i mitten ner medan de andra fortsätter i samma takt. Effekten blir att hela gruppen viker in mot sig själv. På samma sätt böjer sig ljuset inåt, mot mitten, där intensiteten är som störst. Detta fenomen kallas självfokusering, och leder till att pulsen blir ännu mer intensiv. Detta leder i sin tur till mer självfokusering, tills ljuspulsen blir så extremt intensiv att den börjar slita loss elektroner från gasens atomer. De fria elektronerna leder till att laserpulsen slutar fokusera och börjar expandera, varpå intensiteten minskar och processen kommer tillbaka till självfokuseringen, varpå förloppet upprepas. Den sammantagna effekten blir att laserpulsens diameter visserligen hela tiden varierar, men överlag förblir mycket liten.

För att se hur den här laserpulsen rör sig framåt med en rumslig upplösning som ungefär motsvarar dess tjocklek om 30 mikrometer, så krävs det en slutartid om endast 100 femtosekunder (0.000 000 000 000 100 sekunder). Motsvarande bildhastighet är hela 10 tusen miljarder hertz, 10 THz. Mer välbekant är den tusenfaldigt mindre enheten GHz (gigahertz), som ofta används för dator-processorers hastighet (typiskt ~ 3 GHz). En kameras sensor liknar på flera sätt en processor, och har därför en liknande högsta hastighet. Det är alltså omöjligt att filma ens i närheten av tillräckligt snabbt (mer än 1 THz) med en vanlig höghastighetskamera.

Lösningen är att använda korta belysningspulser. Fastän kameran egentligen är öppen under en längre tid, träffar ljuset sensorn bara under de fåtal femtosekunder som belysningspulsen varar. Dessa pulser är ofta runt 30 femtosekunder långa och kan liknas vid extra tunna pannkakor, ungefär tjocka som ett tiondels hårstrå (10 mikrometer). Genom att ljusblixten i sig är så oerhört kort, ”fryses” skeendet trots att sensorn i sig själv är långsam.

Föreställ dig att du tittar ifrån kamerans perspektiv. Från vänster kommer ett laserfilament farande med ljusets hastighet. Samtidigt kommer en pannkaksliknande laserpuls - belysningspulsen - farande rakt mot dig och passerar på sin väg genom filamentet. I ögonblicket när belysningspulsen och filamentet möts, så präglas en bild av filamentet in i belysningspulsen, vilken sedan registreras av kamerans sensor. Vill man ta mer än en bild skapar man ett helt tåg av korta ljuspulser, som passerar filamentet i snabb följd, så att varje puls fångar en ny fas i förloppet.

Här uppstår dock ett praktiskt problem: alla belysningspulser hamnar ovanpå varandra på sensorn, vilket gör det till synes omöjligt att särskilja dem. Lösningen är att varje puls märkts med ett unikt randmönster. Mönstrena kan skilja sig från varandra genom att ha olika frekvens (tjocklek på ränderna) eller olika vinkel. I kamerabilden kommer de överlappande randmönstrena synas och med hjälp av en datoralgoritm kan man i tur och ordning låsa inpå de olika mönstren och återskapa den inpräglade bilden. Det är mycket likt hur man låser inpå en viss radiokanal genom att välja exakt rätt frekvens och därmed filtrera bort alla andra kanaler.

Under mitt doktorandarbete har jag använt denna metod för att filma bland annat laserfilament med mycket hög tids- och rumsupplösning. Jag har dessutom varit med och utvecklat tekniken så att den blivit känsligare för skillnader i brytningsindex, vilket stundtals har varit ett krav för att överhuvudtaget få någon bild. Sammantaget har mitt arbete inneburit ytterligare ett litet steg mot verkligt ultrasnabb videografi, där vi kan avbilda naturens mest flyktiga skeenden och fördjupa vår förståelse av hur ljus och materia samverkar på rekordkorta tidsskalor.

Optical Multiplexing for Ultrafast Videography

I saw that wisdom is better than folly,
just as light is better than darkness.
— King Solomo

Chapter 1

From cave paintings to high-speed videography

1 The importance of observations

Observations lie at the heart of scientific progress, transforming abstract questions into testable hypotheses. In this section, we explore why observations are indispensable, how they help us understand the natural world, and what makes some observations more valuable than others.

1.1 An unprecedented success story

Over the past 300 years, people living in most parts of the world have enjoyed such a great and steady improvement in living conditions, that its like is unprecedented in human history. The kinds of foods that ordinary wage earners can now purchase at a local grocery store, for only a fraction of their income, would have made 17th-century royalty green with envy. Today, we casually buy exotic fruits in the middle of winter and fresh fish in the heat of summer, as though it were the most normal thing in the world.

For much of human history, scarcity, rather than abundance, defined the typical human experience. Most people labored intensively to secure every calorie, enduring the toil of "eating bread in the sweat of one's face," and trading whatever small surplus they could produce for the bare necessities of life. Tragically, it was also common for nearly half of all children to die from disease or injury, with child mortality rates hovering around 50% throughout much of history [1]. Today, by contrast, the global child mortality rate has dropped to approximately 4.3%, and in Sweden, it is now below 0.3%. This remarkable transformation raises an important question: what changed?

A compelling answer is the rise of modern science—often referred to as the scientific revolution—which led to better agricultural methods, vaccines, and an essential understanding of how diseases spread. While this explanation is apt, the term "revolution" can be misleading. In the late 16th century, there was no unified scientific community to be overturned, and the changes that did occur did not constitute a clean break with earlier practices [2]. Instead, a gradual transformation took place, rooted in the principle that reason must be paired with observation and experimentation. Moreover, the development of advanced observational tools—such as telescopes, microscopes, and spectrometers—enabled a broader range of measurable phenomena, pushing the frontiers of science even further.

Pure reasoning, as employed by many ancient Greek philosophers, can prove powerful in mathematics yet falls short in explaining natural phenomena and the laws that govern them. Conversely, observation without accompanying theoretical reasoning—like the extensive recordings made by Babylonian astronomers—can yield valuable data (for instance, determining the best times to sow and harvest) but does not necessarily reveal the deeper nature of our universe. By uniting direct observation with theoretical analysis, modern science was born.

Galileo Galilei personified this approach, demonstrating the importance of empirical observation. Through careful experiments and telescope-based observations, he refuted the prevailing geocentric worldview. Although he faced resistance in his time, his insistence that scientific claims be grounded in observation proved remarkably fruitful. Albert Einstein later referred to Galileo as "the father of modern science," recognizing his role in championing observation as an indispensable element of scientific development [3].

1.2 Underlying assumptions

For modern physicists, the importance of observation and experimentation is self-evident. However, throughout most of human history, it was not generally acknowledged as the foundation of all scientific inquiry. One reason is that making systematic observations of any phenomenon is demanding and thus requires a firm belief that such efforts can yield genuine knowledge. This belief, in turn, rests on several bold and ultimately unprovable assumptions.

The first assumption is that the world around us is "sane" in the sense that it is both real and ordered. This position aligns with classical *realism*, in which an external, mind-independent reality is taken to exist [4, 5]. [1] If the external world existed only in my imagination (or was

---

[1] The references in this section are included primarily to show that similar ideas have been extensively discussed in the literature rather than to provide evidence that the ideas are true. The views presented here

conjured by an evil spirit [6]), I would have no way of knowing otherwise. In that case, it would be absurd to invest time in studying what might be an entirely imaginary reality. Moreover, if the world were not fundamentally consistent in a manner comprehensible to human thought, observations would serve little purpose. If, for instance, the laws of physics changed arbitrarily over time or varied from place to place, observations made today would hold no relevance tomorrow, and any experiment would essentially be a random number generator. This concern is closely related to the "problem of induction," which was explored by David Hume [7]. Hume pointed out that assuming the future will resemble the past in any systematic way cannot be justified purely by logical deduction, emphasizing the leap of faith inherent in our scientific assumptions.

The second assumption is that our senses are trustworthy. This premise is central to *empiricism*, notably championed by John Locke, who argued that knowledge stems primarily from sensory experience [8]. All observations must ultimately be processed through human senses, and we consequently choose to rely on them, even though this reliance cannot be rigorously proven without falling into circular reasoning. Furthermore, our senses occasionally deceive us: we hear or see things that are not truly there, or fail to notice what is plainly in front of us. In the context of this work, my colleagues and I have often stared at image data wondering, "Are we really observing a pattern, or are our eyes overly eager to find structure where there is none?" Despite these occasional pitfalls, we generally trust our senses because, in most cases, they serve us well—and we lack any viable alternative.

The third assumption is that our minds are capable of forming rational conclusions from these observations. This view draws on both *rationalist* traditions—exemplified by René Descartes [6]—and later by Immanuel Kant, who proposed that certain fundamental cognitive structures allow us to meaningfully interpret sensory data [9]. Yet again, trying to prove the rationality of our minds using rational argument leads to the same problem of circular reasoning. If our minds cannot reliably interpret observations, then making those observations would be ultimately pointless.

Given that these assumptions are unprovable, one might ask why any rational person would choose to put so much effort into systematic observation and experimentation. The answer lies in how extraordinarily well this approach have worked. "The proof of the pudding is in the eating," as the saying goes. By following Galilei's principles for only a few centuries, we have not only arrived at a worldview that does not immediately collide with reality, but we have also dramatically reduced deaths from diseases such as tuberculosis and malnutrition. Experiencing such success firsthand provides powerful motivation to continue applying these methods—and makes abandoning them nearly unthinkable.

---

are my own, based on the general stance that observation and experimentation rely on certain philosophical assumptions that cannot be proven but have nevertheless proven practically successful.

1.3 The indirect nature of observations

Even if we accept that observations can be useful, we must still acknowledge that all observations are indirect. In other words, we can never be in direct contact with the object under study; instead, we rely on a series of mediators that convey information from the phenomenon to our minds, where we interpret it.

Consider a red brick wall. At first glance, it seems straightforward to say, “I see a red wall.” In reality, the process is far more complex. External light (for instance, sunlight) interacts with molecules on the brick’s surface, absorbing some wavelengths while reflecting those corresponding to the color red. The reflected light reaches our eyes and triggers responses in our photoreceptors—specifically, one type of cone cell fires more strongly than the other two. These signals are then relayed via nerve impulses to our brains, which ultimately interpret them as “red.”

In this chain of events, our perception depends not only on the physical properties of the brick but also on the spectral distribution of the illumination source, the functioning of our eyes, and how our brains decode the resulting neural signals. If the spectrum of the light were different, or if we were missing one type of cone, our experience of the brick’s color would change dramatically.

Realizing that all observation is mediated can be both discouraging and encouraging. On the discouraging side, it highlights that we can never fully apprehend the object “as it is”; there is always a subjective layer—this “veil of mediation”—between us and the phenomena we study. Yet there is a positive side: even if we for example only manage to capture a partial trace of a phenomenon, [2] we are not fundamentally altering the nature of observation. Since all observations, even the ones that seem most direct, are in fact mediated, relying on incomplete or indirect evidence does not necessarily diminish the value or integrity of our investigation. This realization that every observation is *necessarily mediated* goes back at least to Locke’s distinction between primary and secondary qualities, and remains central to contemporary discussions of perception and knowledge [8].

1.4 What makes an observation “good”?

That all observations are indirect does not mean they are all equally valuable—some are better than others. In scientific practice, the usefulness of an observation is determined by its capacity to address a research question. A visually striking or elegant measurement that does not advance our understanding of a hypothesis has little scientific merit. Conversely,

---

[2] This is precisely the case for several of the experiments presented in this work, where the object under study leaves a “mark” that becomes imprinted on a probe pulse.

even a grainy or low-contrast observation can be invaluable if it sheds light on unexplored phenomena or helps test a theoretical model.

It is also possible for an observation to be scientifically worthwhile when it prompts us to pose new questions rather than immediately answering existing ones. In the context of imaging, this reminder is particularly important. We should not evaluate our methods solely on the aesthetic appeal of the resulting images, but rather on whether they contribute to meaningful insights about the target under investigation.

From a more technical standpoint, an observation becomes far more straightforward to interpret if it responds *monotonically* [3] to some property of the object. In other words, if an increase in the emitted or reflected signal from the target produces a corresponding increase in image brightness, we can more reliably associate changes in pixel intensity with changes in the underlying phenomenon. If the same observed brightness could arise from two fundamentally different states of the target, interpretive ambiguity arises. Ideally, this relationship between a measured quantity and the image signal is not only monotonic but also *linear*, making it simpler to construct accurate models that map measured signals to physical properties.

1.5 Images: a special case of observations

Images are an especially useful category of observations in that they provide a spatially resolved snapshot of a phenomenon or object. Humans rely heavily on vision as our primary, highest-bandwidth sense [10, 11], and we naturally tend to trust what we see. As a result, we often classify and interpret the world according to appearance, which explains why a typical flora for human readers makes extensive use of pictures to identify flowers [4].

Even in the absence of cameras, hand-drawn illustrations can serve as valuable documentation, enabling researchers to compare and discuss specific features with minimal ambiguity. Static images provide a common frame of reference, making it easier to pinpoint critical details and share them with others. Modern cameras have further enhanced this capability, greatly improving both the objectivity and ease of creating images. Tasks that are theoretically feasible without a camera—such as counting stars in the sky or cells in a petri dish—become practically unmanageable without a captured view that can be revisited and methodically analyzed.

Beyond enhancing our ability to preserve and communicate visual information, imaging

---

[3] A function is monotonic if, when the object parameter x changes in a single direction, the observed parameter $f(x)$ does not increase and then decrease, or vice versa. It may, however, remain constant. Mathematically, for a monotonically *increasing* function, $f(x_1) \geq f(x_2)$ if and only if $x_2 \geq x_1$.

[4] A flora for dogs might instead focus on scent, while one for bats would require three-dimensional models of the flower's shape, which would be challenging.

extends the scope of observation into realms far beyond the limits of the naked eye. Cameras can detect light outside the visible spectrum, enabling observations in ultraviolet [12], infrared [13], and X-ray wavelengths [14, 15]. Telescopes collect faint light from galaxies billions of light-years away, allowing astronomers to study galaxies, nebulae, and exoplanets [16, 17], while microscopes provide access to the intricate microcosmos of a single cell, supporting research in biology and medicine, such as diagnosing diseases or studying cellular mechanisms [18, 19]. Aerial and satellite imaging reveal large-scale geological structures, deforestation patterns, or changing coastlines, supporting environmental monitoring and analysis [20]. Time-lapse imaging accelerates slow processes, for example revealing the dynamics and competition involved in plant growth [21, 22], while high-speed cameras capture phenomena unfolding faster than the sense of sight, such as rapid fluid dynamics, shock waves, and laser interactions [23, 24, 25]. Imaging and specifically the ability to capture images has thus revolutionized our perception across spectral, spatial, and temporal dimensions—spanning the microscopic to the cosmic and encompassing events ranging from year-long processes to those occurring in mere milliseconds. This ability to go beyond the natural limitations of human perception underscores the critical role of images in modern scientific exploration.

2 A brief history of images

Far from being merely aesthetic, images stand at the heart of many scientific inquiries. In the following section, the history of imaging will be briefly outlined—from simple hand-drawn depictions to sophisticated high-speed techniques that have allowed scientific discoveries otherwise impossible.

2.1 The earliest man made images

Strictly speaking, an image is any object that represents another thing, generally by resembling it. However, in this work, only flat objects, such as pictures, paintings, or other two-dimensional visualizations that resembles something else, will be considered images. For at least 64,000 years, humans have created objects of this kind, the oldest known being cave paintings of hands (see Fig. 1.1a) made by Neanderthals in what is now modern Spain [26]. These paintings appear to have been crafted by placing one hand (in most cases the left) on a surface and then applying or splashing pigment around it.

Almost 20,000 years later, more figurative paintings began to emerge, often depicting animals. The oldest known example, a painting of a pig found in a cave in modern Indonesia (see Fig. 1.1b) [27], serves to highlight the significant development from the earliest technique—where a physical object leaves an imprint by blocking the application of

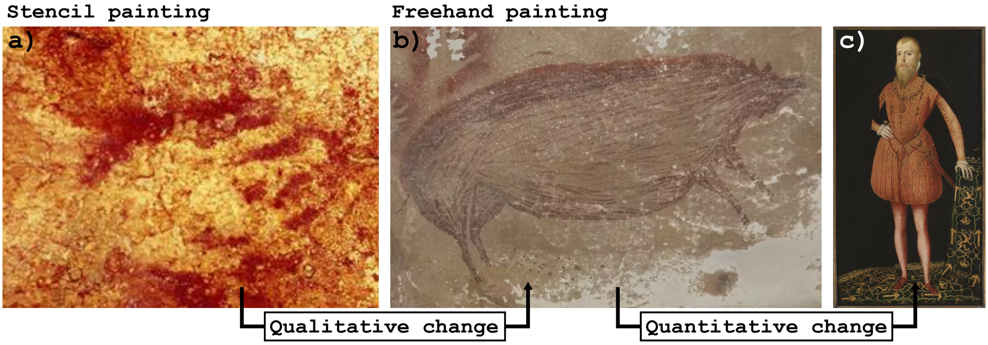


Figure 1.1: Examples of painting techniques. (a) A 64,000-year-old imprint of a Neanderthal hand, created using a stencil painting technique, where pigment was applied around the object [26]. Credit: Hipolito Collado Giraldo. (b) A 45,000-year-old freehand drawing of a pig from Sulawesi, Indonesia [27]. Credit: Maxime Aubert. (c) A painting of Erik XIV from 1561 by Steven van der Meulen. The transition from (a) to (b) represents a qualitative change in technique, while the progression from (b) to (c) is primarily quantitative.

pigment—to freehand drawing from the artist's memory. This transition enabled the creation of images representing not just physical objects but also imagined ones, expanding the scope of what could be visualized.

The further development of painting techniques eventually allowed Dutch painter Steven van der Meulen to paint the realistic portrait of Erik XIV in Fig. 1.1c. This portrait was sent to Queen Elizabeth I of England as part of Erik's ultimately unavailing marriage proposal. [5] The imaging development, however, was not qualitative in the same sense as the earlier shift from stenciling to freehand drawing. To clarify, a qualitative development introduces entirely new capabilities, while a quantitative development improves on existing capabilities. [6] Hence, the shift from rock paintings of pigs in Spain to the detailed portrait of Erik XIV represents a quantitative improvement in an existing technique, even though the progression was undeniably immense.

2.2 Painting with light

The next truly major leap in the development of imaging techniques occurred when images of the environment could be automatically generated by the use of a camera. This breakthrough required two key technical advances. The first was optical imaging, which will just be called ”imaging” from here on. Its essence is that light originating from a particular point on an object should, after the imaging process, converge to a single point on the screen or detector. Additionally, light from neighboring points on the object should result in neighboring points on the screen. If these criteria are met, the entire object—composed

---

[5] To all of Sweden's surprise the Queen of England rejected the gracious offer, but is said to have expressed delight in the noble Swede's slender legs [28, p.406]

[6] A quantitative improvement is thus an improvement of quality, which can be confusing.

of many such points—will appear as an image on the screen. In practice, this projection is never perfect, as light from a point is always spread over a small area rather than converging to a single point. However, the smaller this projected area, the better the imaging quality.

Even if individual rays of light travel in this point-to-point manner, the collection of rays originating (or reflected) from a given point on an object spreads in all directions. This phenomenon (fortunately) allows objects to be seen from different angles but also (unfortunately) means that imaging does not occur by itself. Achieving imaging therefore requires some ingenuity. The simplest and historically first solution relied on the realization that some rays of light do follow the desired paths that allow imaging. The task is just to filter out all the other rays.

This can be achieved by placing the screen in a completely dark room and making a tiny hole in the wall opposite it. In this configuration, all light reaching the screen must pass through the pinhole. Rays constrained by both their point of origin and a point along their propagation will travel along a single line and converge at the corresponding point on the image screen. This setup satisfies the imaging criterion, as illustrated in Fig. 1.2a, which shows how light cones from a traffic light form an inverted image inside the camera obscura. The traffic light (which is obviously broken, as all three colors are lit simultaneously) emits light evenly in all directions, but for clarity, the rays that pass through the pinhole and eventually form the image have been enhanced.

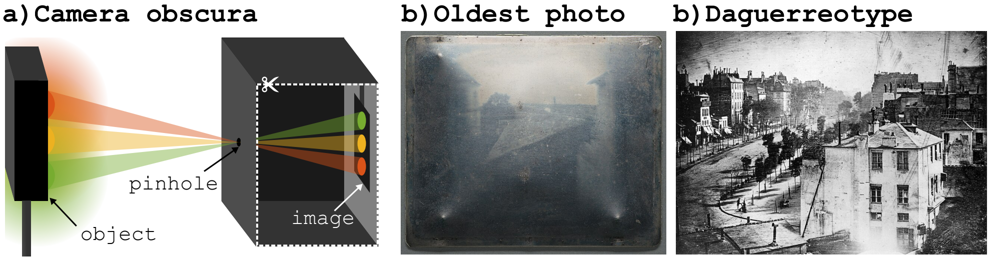


Figure 1.2: Milestones in early photographic development. (a) Illustration of the camera obscura principle using a traffic light as the object. The light cones from each lamp that pass through the small pinhole (artificially emphasized relative to the filtered-out light) project an inverted image onto the screen inside the dark room. The dashed line indicates a cutaway side wall, removed for demonstration purposes. (b) The oldest known surviving photograph, created around 1827 by Joseph Nicéphore Niépce, using a bitumen-coated plate and a camera obscura. (c) One of the earliest daguerreotypes, depicting a view of the Boulevard du Temple in Paris, created by Louis Daguerre around 1837.

The name of this technique, *camera obscura*, has given us the word ”camera,” but literally means ”dark room” [29]. Its first documented mention in literature dates back to 400 B.C. in China [30]. While this invention represented a significant step forward in imaging, it had practical limitations: the tiny pinhole allowed only a small amount of light to reach the screen, necessitating long exposure times.

To address this limitation, lenses were introduced, allowing more light to be collected while still maintaining the imaging properties of the pinhole. By focusing light rays from a single point on the object to a single point on the screen, lenses preserved the point-to-point correspondence essential for imaging. This innovation drastically reduced exposure times and paved the way for modern cameras. However, for the formed images to be permanent (as in painting techniques), another technical solution was needed.

The second technical advancement was the development of photosensitive materials, which chemically change under the influence of light and gradually darken. Areas exposed to more light appear darker in the developed image, effectively inverting the grayscale. This principle was first demonstrated in 1827 by Joseph Nicéphore Niépce, who used a bitumen-coated metal plate as the screen in a camera obscura to record an image. The oldest surviving image created in this way is *View from the Window at Le Gras*, shown in Fig. 1.2b. Although a visual image would form, it faded soon after development [31]. Despite this limitation, Niépce's experiment marked a milestone: light, rather than pencils, had been used to draw an image.

In 1839, Louis Daguerre presented to the world a technique that allowed the creation of permanent images through light exposure—the daguerreotype. Producing a daguerreotype involved polishing a silver plate to a mirror finish, rendering it photosensitive with iodine vapor, developing the photo using mercury vapor, and desensitizing the plate with a hot salt solution. Despite the effort involved and the fragility of the result—the photo could be destroyed by the slightest touch—the daguerreotype was a success and dominated photography for two decades [32]. An early example, showing a Parisian boulevard is shown in Fig. 1.2c. Due to the long exposure time, moving traffic is absent from the image, but two stationary figures are visible, making the photograph widely regarded as the oldest surviving photo of humans. Can you spot them? [7]

Since then, much more practical photosensitive materials have been developed. However, the basic process of creating a photograph—using light to chemically alter a material—remained largely unchanged until the advent of Charge-Coupled Devices (CCDs) in the late 1960s [33] and the first commercial digital camera in 1988 [34].

A common characteristic of all the types of images described so far is that they exist only as physical objects. This means the image is inseparably tied to the material from which it is constructed—whether pigment on rock, oil on canvas, or bitumen on a copper plate. If we attempt to deconstruct the image into its visual essence and store it purely as an idea, it is inevitably lost. Even if the image persists in the mind of the most skilled painter, it can never be perfectly recreated in its original form; each attempt will result in a slightly altered rendition. Consequently, the full essence of such images can exist in only one representation: the actual physical image itself.

While information can be extracted from these images—such as the physical size or ar-

---

[7] Located in the center of the lower-left quadrant, one person is polishing the other's shoes.

rangement of the objects depicted—this information is always incomplete and cannot fully convey every aspect of the image. Thus, for all images up to this point in history, the image itself is the primary entity, while the information it contains is secondary.

2.3 Digital images

This fundamentally changed in 1957, when Russel Kirch at the National Institute of Standards and Technology scanned a photograph of his son (see Fig. 1.3a), creating the first digital image [35]. The image only measured 176 by 176 pixels, but it was the start of a new imaging era. In digital images, information takes precedence, while the actually visible image comes second and can be said to be a mere realization of the preexisting image information. In short, the essence precedes the existence. [8] This breakthrough allowed images to exist beyond their physical manifestation and enabled entirely new possibilities: images could be transmitted, manipulated, analyzed, and transformed in ways that were previously inconceivable. The ability to transform and represent the exact same image information in completely different forms is central to the work presented in this thesis.

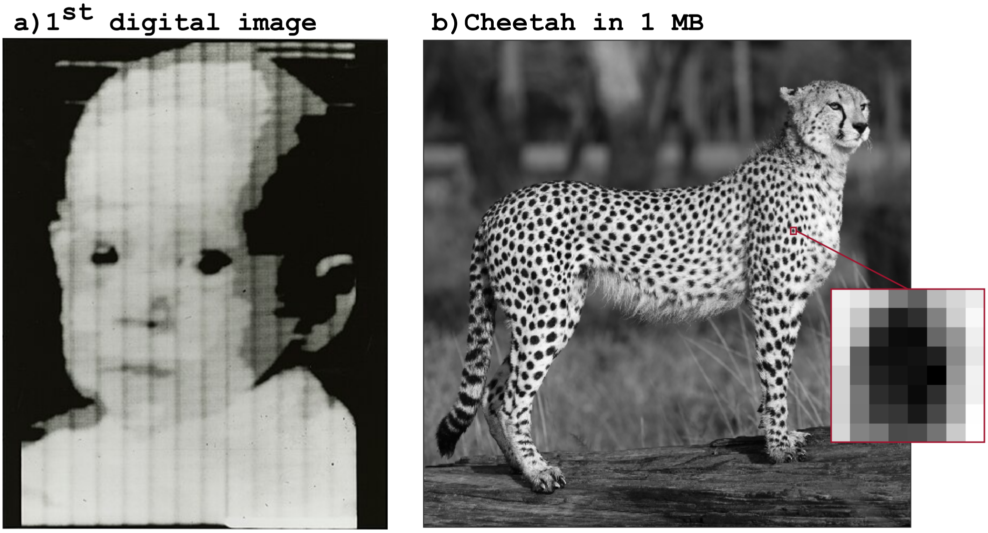


Figure 1.3: Examples of digital images. (a) The first-ever recorded digital image, depicting the son of its creator, originally scanned at a resolution of 176 by 176 pixels. Public domain [37]. (b) A digital image of a cheetah with a resolution of 1000 by 1000 pixels and 256 intensity levels, corresponding to 1 MB of data. The zoomed-in section reveals the discrete pixel structure of the image. Credit: Tara Turkington, Flickr: flowcomm.

---

[8] Now all children reading this think: You forget about the fuse bead plate! Doesn't the essence precede the existence for fuse beads with pre-made patterns? You are right. However, Gunnar Knutsson did not invent fuse beads until 1962 [36], five years after the first computer image. Still, the fuse bead plate is a type of digital image, since it is discretized both in space and color selection. And so are knitting patterns. And carpet patterns. In fact it is quite possible to argue that digital images are very old.

Before delving into these transformative aspects of digital images, it is necessary to first define what a digital image is. A digital image consists of a finite number of elements, each represented by a finite value. These elements may be vectors (as in vector graphics) or mathematical functions, but here we will focus exclusively on raster images—images composed of pixels arranged in equidistant rows and columns. Furthermore, we will constrain our discussion to monochromatic (grayscale) images, as they are more fundamental than polychromatic images (typically represented in red-green-blue channels) and are the primary type used in ultrafast imaging.

In a monochromatic image, each pixel is assigned a value, known as the pixel value, which determines its grayscale intensity. Typically, this value is an integer between 0 and $2^b - 1$, allowing for 2^b possible grayscale levels. Here, b represents the bit depth of the image, indicating the number of *bits* required to store each pixel value. Common bit depths include 8, 10, 12, and 16, where a bit depth of 8 or 16 corresponds to 1 or 2 *bytes* (B) per pixel, respectively. For example, the image in Fig. 1.3b is 1000 by 1000 pixels, with a bit depth of 8 ($256 = 2^8$ grayscale levels), requiring precisely $1000 \cdot 1000 \cdot 1$ byte $= 1$ MB of storage.

Unlike analog images, the data required to store a digital image is finite and clearly defined. The trade-off for this precise quantization is the pixelated nature of the image, as demonstrated by the zoomed-in section of one of the cheetah's spots in Fig. 1.3b. However, as long as we do not magnify the image excessively, the limiting factor for observable detail will not be the image data itself but rather the resolution of the printing process—or, ultimately, the resolution of our eyes. When the resolution of the human eye is the limiting factor, we perceive the image as continuous, despite its inherently discrete nature.

Since image information in digital form is fully accessible, it can be transformed in ways that make otherwise undetectable features visible. This capability alone is such a tremendous benefit that it justifies sacrificing most [9] of the inaccessible analog image information to obtain the accessible digital data. Transforms allow the representation of an image to change: the resulting representation might become unrecognizable as the original image or recorded object, or it might remain recognizable but altered in some way. Mathematically, a transform is a function that applies a set of rules to a dataset, mapping it to a new representation. In the case of lossless transforms, this new representation retains the same information as the original, enabling perfect reconstruction. By contrast, lossy transforms intentionally discard certain information to simplify or compress the data. Two notable examples of lossless transforms are the Fourier Transform and the Wavelet Packet Transform.

Figures 1.4a and 1.4b illustrate these transforms using the cheetah image from Fig. 1.3b. In the Fourier-transformed image (Fig. 1.4a), the cheetah becomes entirely unrecognizable, while in the wavelet-decomposed image (Fig. 1.4b), the cheetah is more or less recognizable,

---

[9] In fact, almost all information is discarded during the digitization of an analog image, as the difference between infinite and finite is vast.

depending on which sub-image you look at.

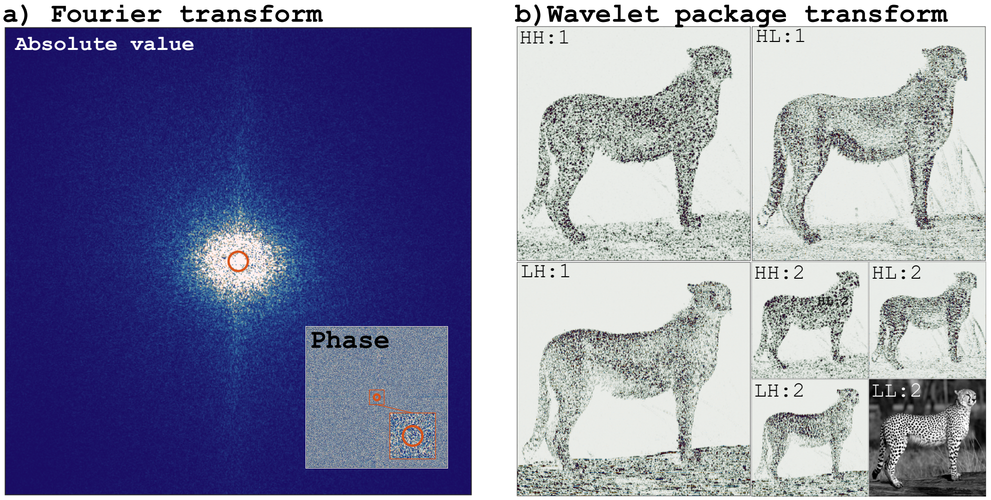


Figure 1.4: Examples of lossless digital image transforms applied to the cheetah image from Fig. 1.3b. (a) Fourier transform. The large image shows the absolute value of the image's frequency components, with larger values concentrated near the center (low frequencies). The small image displays the phase of the frequency components, which appears random but is essential for reconstructing the image. (b) Wavelet packet transform at two levels. HH: diagonal edges; HL: horizontal edges; LH: vertical edges; LL: low-resolution approximation.

The Fourier Transform represents the image in terms of spatial frequencies—essentially wave patterns of varying wavelengths and orientations—that must be combined to reconstruct the original image. An intuitive analogy is a skilled musician identifying the individual notes in a chord; probably without realizing it, she performs a Fourier Transform in her mind by isolating the sound's frequency components. Consequently, the Fourier Transform is often referred to as a frequency transform, and the transformed data is said to exist in the Fourier or frequency domain.

For each pixel in the original image, the Fourier representation contains two numbers: an absolute value, which indicates the strength of each frequency component, and a phase, which determines the component's alignment. The inset in Fig. 1.4a shows what appears to be random noise but actually encodes the critical phase information required to accurately reconstruct the original image. Lower-frequency components are typically shown closer to the center of the Fourier transform, while higher-frequency components are positioned further out. Although the Fourier transform appears to require two numbers (amplitude and phase) for each pixel in the original image, its inherent symmetry ensures that the total amount of information is conserved between the original image and its transformed representation. [10]

---

[10] In the Fourier domain, amplitude and phase jointly encode the information of the original image. Low-frequency components typically have large amplitudes, where minor phase variations have little effect, while high-frequency components have small amplitudes but require precise phase information for accurate recon-

The Wavelet Packet Transform, on the other hand, decomposes an image into four sub-images, each with one-quarter of the original pixel count. These sub-images capture different types of information: diagonal edges (HH [11]), horizontal edges (HL), vertical edges (LH), and a low-resolution approximation of the original image. In the example shown in Fig. 1.4b, the first-level approximation is not explicitly presented but has been further decomposed into its own components. While much more could be said about these transforms, this brief overview demonstrates how digital images enable representations of image information that are otherwise impossible.

Beyond being transformable, digital images can be modified and analyzed in ways that are simply not possible for analog images. Figure 1.5 shows three examples of modifications made to the cheetah image from Fig. 1.3b. In the Fourier Transform in Fig. 1.4a, a thin circular red line is visible. By setting all components outside this circle to zero and performing an inverse Fourier Transform, the *low-pass* filtered image in Fig. 1.5a is produced. This image retains only the low-frequency components, resulting in a blurred version of the original image where only the cheetah's general shape is preserved, but its spots are removed.

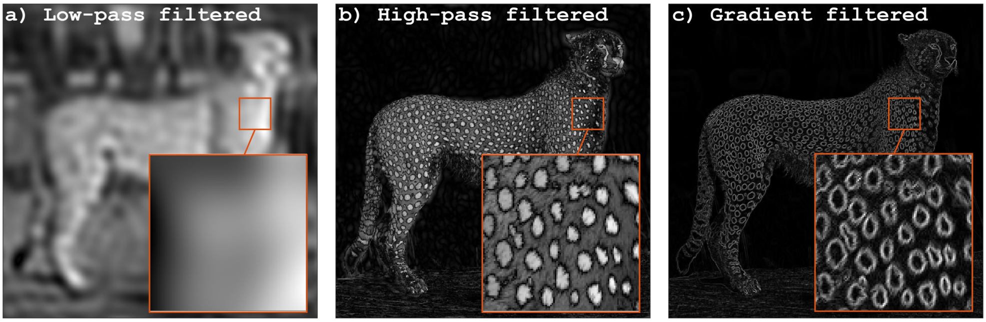


Figure 1.5: Examples of filters applied to digital images, demonstrated using the cheetah image from Fig. 1.3b. (a) Low-pass filtering, performed by Fourier transforming the image, setting the components *outside* the orange circle in Fig. 1.4a to zero, and then inversely transforming back to the spatial domain. (b) High-pass filtering, performed identically to (a), except the components *inside* the orange circle are set to zero. (c) Gradient filtering, performed by pixel-wise calculation of the combined amplitude of the image derivatives in the horizontal and vertical directions.

Where did the spots go? They appear in the complementary *high-pass* filtered image in Fig. 1.5b, which was generated by only keeping the components outside the circle in Fig. 1.4a, and setting those inside to zero. High-pass filtering highlights the high-frequency compo-

---

struction. This interplay suggests a trade-off: components with high amplitude can tolerate lower phase precision, while those with low amplitude demand higher phase precision. For the DC component (representing the average intensity), phase information is entirely irrelevant. On average, the combined bit depth allocated to an amplitude-phase pair must equal the bit depth per pixel in the original image to ensure information conservation.

[11] H and L stand for high- and low-frequency components, respectively. The first and second letters refer to the horizontal and vertical directions, respectively. HH hence represents high-frequency components (i.e., edges) in both directions, meaning diagonal edges.

nents, such as the spots, making them more prominent. By applying specific types of filters—low-pass, high-pass, or even band-pass filters (which isolate a selected frequency band)—different aspects of an image can be emphasized or suppressed, facilitating detection and analysis.

Finally, in Fig. 1.5c, the gradient of the original cheetah image has been calculated, accentuating transitions between regions and aiding in tasks such as edge detection. These examples are just to show a tiny fraction of what can be done with digital images, in terms of modification and analysis.

2.4 Recording digital images

The following section introduces the devices used to record digital images, beginning with the earliest one.

2.4.1 The birth of digital image sensors: the charge-coupled device

To automatically produce digital images using a camera or other imaging system, photosensitive film had to be replaced by a device capable of quantifying light intensity at distinct locations. The Charge-Coupled Device (CCD), invented in 1970, achieved this goal and had a revolutionary impact on imaging [33]. It consists of a pixel array, with one photoactive region per pixel, which generate charge upon exposure to light. Connected to the photoactive regions are potential wells, which can collect and transfer charge. The architecture is sketched in Fig. 1.6a.

The photoactive region is made of a doped semiconductor with an external static voltage applied, creating a depletion zone devoid of free charge. During exposure, photons striking atoms in the depletion zone generate electron-hole pairs. The external voltage accelerates the free charges, directing the electrons into the potential wells. The amount of charge collected in each well corresponds to the intensity of light at that pixel.

After exposure, the collected charge must be digitized. This process involves sequentially shifting the charges from all pixels toward a single Charge Voltage (CV) converter, which amplifies the charge into an analog voltage. The analog voltage is then digitized by an Analog-to-Digital (AD) converter. The readout begins with the lowest row of pixels, where each pixel's charge is digitized one at a time. Once the first row is fully processed, the remaining rows are shifted down by one step, and the process repeats. This sequential transfer of charge operates similarly to a bucket-brigade system, where each pixel contains three potential wells. By modulating the depth (voltage) of these wells in a controlled sequence, the charge ”pours” from one well to the next, ultimately reaching the CV and AD

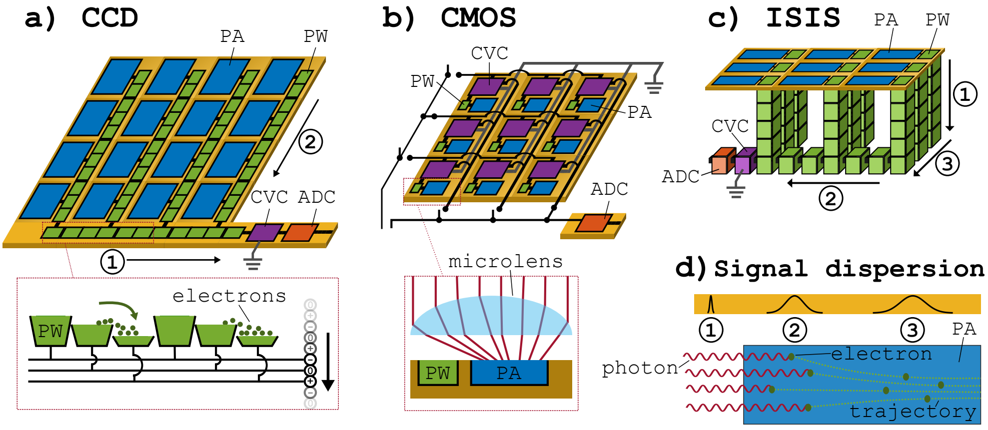


Figure 1.6: Architectural evolution of digital image sensors. (a) The Charge-Coupled Device (CCD) architecture relies on charge generation in a photoactive area (PA), sequential charge transfer to a single Charge Voltage Converter (CVC) that amplifies the signal, and digitization by an Analog-to-Digital converter (ADC). (b) The Complementary Metal-Oxide-Semiconductor (CMOS) architecture has one CVC per pixel for parallel readout, enabling higher frame rates. (c) The In-situ Storage Image Sensor (ISIS) architecture allows ultra-high-speed burst imaging by temporarily storing charge in the vertical stack of potential wells (PW) within the pixel array. (d) The ultimate frame rate limit for silicon-based sensors is determined by the temporal signal dispersion, caused by randomness in photon-electron conversion and electron drift, yielding a theoretical maximum of 90.1 Gfps for green light.

converters without mixing with the charges from adjacent pixels. This method, illustrated in the zoomed in section of Fig. 1.6a, ensures accurate digitization of the accumulated charge into digital values.

One key advantage of CCD sensors is that all pixels share the same converters, ensuring that any systematic errors introduced during digitization are consistent across the entire detector. However, this architecture also imposes limitations on read-out speed, making high frame rates difficult to achieve. To increase the frame rate, the rate of charge shifting must be accelerated, which reduces the efficiency of charge transfer between wells and increases noise levels.

Despite their limitations, the highest-performing CCD sensors can achieve frame rates of approximately 10 fps for large-format sensors (e.g., 16 MP) and significantly higher for smaller sensors. These sensors also excel in spatial resolution, reaching over 100 MP [38]. Additionally, their shared CV and AD converter architecture ensures exceptionally low noise levels, particularly when cooled to reduce thermal noise. This combination of high resolution and low noise makes CCDs ideal for applications requiring maximum sensitivity. Consequently, they remain widely used in fields such as astronomy [39], where long exposure times and the highest possible signal-to-noise ratios are essential.

2.4.2 Revolutionizing speed: the complementary metal-oxide-semiconductor.

Similar to CCD sensors, Complementary Metal-Oxide-Semiconductor (CMOS) sensors consist of a pixel array, with each pixel containing a photosensitive region and a charge storage area, as illustrated in Fig. 1.6b. However, unlike CCDs, CMOS sensors integrate a CV converter and additional circuitry directly within each pixel. This circuitry connects the pixel to an AD converter via electronically controlled switches. By selectively opening and closing these switches, each pixel can be independently read out, eliminating the time-consuming charge transfer process required in CCDs. This design allows for significantly faster readout of the entire array, as voltage signals propagate much faster than physical charge transfer [40].

This architecture, however, introduces two notable drawbacks. First, the CV converters are not perfectly identical, leading to fixed-pattern noise (FPN), which manifests as slight variations in the digitization process. Second, the CV converters and associated circuitry occupy part of each pixel, reducing the photosensitive area and overall sensitivity. Both challenges can be mitigated. Fixed-pattern noise is addressed by calibrating each CV converter using known quantities of charge, ensuring consistent output across the array. Sensitivity loss is mitigated by integrating micro-lenses atop each pixel, which focus incoming light onto the active photosensitive regions. While micro-lenses are commonly associated with CMOS sensors, they are also employed in advanced CCD designs to enhance light-gathering efficiency. CMOS sensors that incorporate pixel-wise CV calibration are referred to as scientific CMOS (sCMOS) sensors, specifically tailored for scientific applications. Today, sCMOS sensors have largely replaced CCD sensors in most scientific contexts due to their superior speed, lower power consumption, and comparable or even better noise performance. [12]

With an individual CV converter for each pixel, CMOS sensors eliminate the bottleneck associated with the sequential charge transfer process in CCDs. Instead, the primary limitation for achieving higher frame rates shifts to the time required to digitize and store data from each pixel. Frame rates can thus be further increased by reducing the number of active pixels, either through pixel binning or by selecting a smaller region of interest. For instance, commercially available high-speed CMOS cameras, such as the Photron E9 100S, achieve frame rates of 326 thousand frames per second (kfps) ($\Delta T = 3.1$ µs) with a resolution of 640 × 480 pixels. By reducing the resolution to 640 × 32 pixels, the frame rate increases to 2.72 million frames per second (Mfps) ($\Delta T = 368$ ns). This capability to trade field of view for temporal resolution makes CMOS cameras exceptionally versatile, allowing them to capture phenomena across a wide range of timescales, including ballistics

---

[12] During my work, I received funding to purchase a high-resolution CCD camera. However, by the time I was ready to order, the manufacturer strongly recommended switching to an sCMOS camera. Although the CCD camera was still in stock, the sensor itself was no longer in production, meaning repairs or replacements would not be possible if the sensor failed.

[41], fluid dynamics [25], and combustion processes [23].

2.4.3 A faster architecture: the in-situ storage image sensor

In 1996, Kosonocky et al. introduced a groundbreaking approach to overcome the data storage bottleneck associated with CMOS sensors [42]. Their solution involved surrounding each pixel with additional, sequentially ordered storage elements. During acquisition, photogenerated charge is continuously and sequentially transferred from one storage element to the next in what can be described as a "burst". This transfer within each pixel's own "shift-register" is represented by arrow 1 in Fig. 1.6c. In modern designs, the shift registers are typically stacked beneath the primary photosensitive region of the pixel, maximizing the effective light-collection area while maintaining the compact architecture. Only after the full acquisition sequence is completed is the accumulated charge digitized and stored to memory. This approach eliminates the need for high-speed data transfer to external memory during acquisition, drastically increasing achievable frame rates. The initial prototype utilized 30 storage elements per pixel, thus enabling a sequence depth of 30 frames. The technique was originally referred to as a Very High Frame Rate (VHFR) burst image sensor but later became widely known as the In-situ Storage Image Sensor (ISIS) [43].

Over the years, the ISIS architecture has undergone significant advancements. By optimizing the geometry to minimize abrupt changes in electron movement, the technology has evolved dramatically. The original design achieved frame rates of 0.5 Mfps ($\Delta T = 2$ µs) across 30 frames with a pixel resolution of 360 × 360 pixels. Recent iterations, such as the version described by Suzuki et al., have reached frame rates of 125 Mfps ($\Delta T = 8$ ns) across 368 frames, with a resolution of 50 × 108 pixels [44].

A constraint inherent to this type of sensor is the fixed sequence depth, determined by the architecture. Additionally, there is a trade-off between total pixel count and frame rate, which arises from geometric constraints in the design. As a result, achieving the highest frame rates often requires a reduction in pixel resolution. Despite these limitations, ISIS cameras remain a highly effective solution when CMOS cameras are too slow. They have been successfully applied in both fundamental [45] and applied research [46, 47], demonstrating their utility across a range of scientific and engineering disciplines.

2.4.4 Pushing limits: the ultimate frame rate of silicon sensors

The ultimate frame rate achievable with silicon-based sensors, including CCD, CMOS, and ISIS, is fundamentally constrained by the physics of the light-matter interaction within the silicon photoactive region. At the core of this limitation lies the temporal broadening of the signal during photon-to-electron conversion, which eventually prevents temporally distinct events from being resolved because of charge mixing. Specifically, the depth at which incoming light interacts with silicon atoms to generate electron-hole pairs is inherently random. For green light, the average penetration depth is approximately 1.73 µm, but individual photon interactions can occur at shallower or deeper locations. This randomness causes the initial positions of the generated charges to be distributed within the silicon. Once generated, these charges start to drift toward the potential well. However, because of their finite velocities and varying initial positions, the times they take to reach the well are also randomly distributed. This temporal dispersion leads to mixing of charges associated with distinct optical events, making it increasingly difficult to resolve events that occur close together in time. The stochastic nature of the light-matter interaction and subsequent random electron drift thereby imposes a lower limit on temporal resolution, which translates to an upper limit for the achievable frame rate.

This physical limitation is illustrated in Fig. 1.6d, which schematically outlines three steps in the process and the associated temporal signal dispersion: (1) photon propagation through air and silicon, (2) the generation of photoelectrons at random depths, and (3) the subsequent drift of electrons towards the potential well. Simulations suggest that the theoretical maximum frame rate for silicon-based sensors is 90.1 Gfps for green light. Beyond this threshold, the intrinsic properties of silicon prevent further improvements in temporal resolution, establishing an absolute boundary for these systems [48].

Chapter summary

Modern science, one of humanity's greatest success stories, is founded on experimentation and observation. Among the most transformative forms of observation are images, which not only enhance the communication of scientific findings but also expand the temporal, spatial, and spectral range of human vision. The advent of digital imaging has unlocked unprecedented possibilities for analyzing and modifying images, as well as capturing events at extremely high speeds. However, digital sensors face an inherent upper limit in achievable frame rates. To push beyond this boundary and capture even faster phenomena, fundamentally new and qualitatively different imaging approaches must be developed.

Chapter 2

Ultrafast videography

1 Why bother going faster?

The last chapter concluded that digital sensors face an insurmountable upper limit of 90 Gfps, corresponding to an inter-frame time of approximately 11 ps. This limit is imposed by the fundamental physics of silicon-based sensors, and no amount of quantitative improvement can overcome it. If we want faster image sequences a qualitatively different approach is definitely needed, but the question remains: is there any real need to film faster than 90 Gfps?

1.1 Curiosity drives us

As humans, we naturally inhabit the realm of meters and seconds—scales that align with the objects and timescales we encounter daily. While we frequently work with units like kilometers (1,000 meters) and millimeters (1/1,000 meter), distances beyond these—like megameters (1,000 kilometers) [1] or micrometers (1/1,000 millimeter)—become almost incomprehensible without the aid of specialized tools. Instruments such as maps, bicycles and microscopes bridge these gaps, transforming the inaccessible into something tangible. Similarly, motion that unfolds over durations longer than 1,000 seconds or shorter than 1/10 second can be very difficult to perceive, as it lies outside our natural ability to follow events that are either too slow or too fast for our senses.

Historically, humanity has pushed beyond these limits, exploring both the vast macrocosmos, where distances are measured in light-years and time in millions of years, and the intricate microcosmos of the atom, where distances are measured in nanometers (10^{-9}

---

[1] Yes, you can probably imagine 1 megameter, but I am pretty sure you do it on a map and not in ”real” life.

m) and time in attoseconds (10^{-18} s). In doing so, we have made remarkable discoveries. Some, like gravitational waves [49] and exoplanets [50] in the macrocosmos, or the electron [51] and DNA [52] in the microcosmos, were predicted long before their eventual confirmation. Others were, however, entirely unexpected, arising from chance observations or experimental surprises, such as the cosmic microwave background [53] and the existence of pulsars [54] in the macrocosmos, and quasicrystals [55] and bacteria [56] in the microcosmos.

The historic record of science thus teaches us that discovery is not always the result of a targeted effort; many breakthroughs occur unexpectedly. However, such chance discoveries often arise when we actively strive to extend the boundaries of observation and measurement. This is why we must continue advancing imaging and videography technologies, pushing beyond the limits of silicon sensors and exploring the ultrafast regime beyond 1 Tfps (10^{12} frames per second). Even if we do not know precisely what we are looking for, such efforts may reveal phenomena we cannot yet conceive—discoveries that no one has predicted. To claim that everything is already known, or that extending videography into the ultrafast domain is futile, would not only be shortsighted but also contrary to the very spirit of scientific inquiry.

Curiosity must therefore drive us, even if no other motivation existed. Fortunately, we are not searching blindly; there are compelling reasons to push these boundaries, due to the existence of scientifically interesting ultrafast phenomena.

1.2 Potential ultrafast targets

In the following section, I will explore a selection of phenomena that ultrafast imaging has the potential to illuminate, offering new insights into the dynamics of the natural world.

1.2.1 Laser wakefield acceleration

Laser Wakefield Acceleration (LWFA) is a technique for accelerating electrons using ultrashort laser pulses instead of conventional external electric fields, as in linear or circular accelerators. This method achieves accelerating fields on the order of GeV/m, far surpassing the MeV/m fields of traditional accelerators. Consequently, LWFA can produce electrons with equivalent energies over distances that are three orders of magnitude shorter [57]. These high-energy electrons have applications in medical imaging and therapy, materials science, high-energy physics, and the generation of ultrashort X-ray pulses [58]. While LWFA performance approaches that of classical accelerators, it has not yet fully matched their capabilities [59].

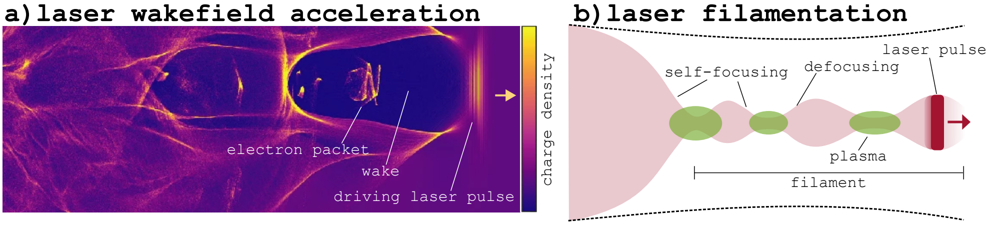


Figure 2.1: Potential targets for ultrafast videography. (a) A simulation of laser wakefield acceleration, propagating to the right. Adapted from [59]. (b) The process of alternating self-focusing and plasma-driven defocusing, which is thought to yield laser filamentation. Adapted from [61].

The accelerating wakefield in LWFA can be likened to a wave that electrons ”surf” on, gaining energy as they ride it [60, 58]. This wakefield is generated when an ultrashort laser pulse interacts with a gaseous medium, ionizing it to create plasma. The plasma, composed of heavy positive ions and lightweight negative electrons, is macroscopically neutral due to the even distribution of charges. As the laser pulse propagates, its transverse (perpendicular) electric field displaces the electrons, pushing them away from the laser’s path while leaving the heavier ions nearly stationary. This displacement creates a positively charged region, or ”wake,” that closely trails the laser pulse. A simulation of this process can be found in Fig. 2.1a.

Electrons, whether randomly present or intentionally injected into this wakefield, can become trapped and accelerated. As they are continuously drawn toward the point of maximum positive charge, they experience not only forward acceleration but also transverse and longitudinal focusing. This process is analogous to a boat moving through water: as the boat displaces water, it leaves a trailing ”hole” bordered by a wave. Any object with a velocity similar to the boat can become trapped in the hole and propelled forward by the wave. Similarly, in LWFA, electrons with appropriate velocities entering the wake are captured and accelerated.

Despite its elegance, LWFA comes with several challenges. One significant issue is dephasing, where electrons outrun the accelerating phase of the wakefield, leading to dispersion and loss. This can be mitigated through spatio-temporal pulse shaping, which creates a phase that travels at the speed of light, preventing electrons from overtaking it [62]. Another complexity arises from the highly non-linear interaction between the laser pulse and the plasma, where effects like self-focusing due to the optical Kerr effect alter the laser field and the resulting wake. Understanding these dynamic processes is crucial for optimizing the acceleration mechanism.

Efforts to monitor LWFA dynamics have employed various imaging techniques. Holography has been used to visualize electron density variations, demonstrating the high sensitivity of interferometric methods, but these were static images and could not capture the temporal evolution of the process [58]. Shadowgraphy has also been employed [62, 60], but also only for single-frame snapshots. To better understand the acceleration process and determine critical experimental parameters, a videographic imaging technique with high sensitivity to electron density variations, spatial resolution on the order of a few micrometers, and matching temporal resolution (20 fs) would be highly beneficial.

1.2.2 Laser filamentation

Laser Filamentation, first reported in 1995 [63], is a phenomenon where a high-intensity, ultrashort laser pulse undergoes self-channeling, forming a thin filament approximately 100 µm in diameter that can extend over at least a kilometer [64]. These filaments have been proposed for applications such as remote sensing of gases and aerosols in the atmosphere [65], lightning control [66], laser-based telecommunications in foggy conditions [67], and laser-induced spectroscopy [68]. Additionally, filamentation plays a crucial role in laser wakefield acceleration (LWFA), as the driving laser pulse often undergoes filamentation. Understanding the filamentation process is therefore essential for developing stable LWFA conditions [69].

Laser filaments form due to a dynamic balance between two competing effects, as illustrated in Fig. 2.1b. The first is self-focusing caused by the optical Kerr effect, where the refractive index n of a material depends on the intensity I of the incident light, such that $n = n_0 + n_2 I$. Since the nonlinear index n_2 is much smaller than the linear refractive index n_0, the Kerr effect becomes significant only at very high intensities, such as those found in ultrashort laser pulses. The typical intensity profile of a laser pulse has a higher intensity at the center and lower intensity towards the edges. Combined with the Kerr effect, this results in a higher refractive index near the center of the beam, creating an optically longer path for the light in the center. This acts like a focusing lens, causing self-focusing if the laser power exceeds a critical threshold P_c.

The second effect, which counteracts self-focusing, is defocusing. This is typically attributed to plasma generation caused by the high-intensity laser light [65]. However, some studies suggest that defocusing may also arise from higher-order nonlinear Kerr effects, such as contributions from n_4, n_6, and beyond [70]. Regardless of the exact mechanism, defocusing eventually counterbalances self-focusing, bringing the pulse back to a Kerr-effect-dominated regime, where it refocuses. This cyclical interplay between self-focusing and defocusing confines the pulse within a small, stable volume, forming a filament.

Laser filamentation remains incompletely understood, with competing models proposing different mechanisms to explain its behavior. To determine which model better describes the process, improved data on the dynamic behavior of filaments is needed. Since the process evolves rapidly, this data should ideally be temporally resolved. As a highly nonlinear and dynamic phenomenon, laser filamentation is particularly sensitive to the initial conditions of the driving laser pulse, making the process inherently stochastic [71]. While this stochasticity may not be apparent on a macroscopic scale, it becomes evident when the process is studied in a microscopic configuration, as demonstrated in Paper v of this thesis.

The stochastic nature of filamentation poses significant challenges for sequential imaging techniques, as they cannot accurately capture the true dynamics of the process due to the variability between events. Instead, the study of laser filamentation calls for single-shot ultrafast videography capable of high spatio-temporal resolution and interferometric sensitivity. These requirements are similar to those for studying laser wakefield acceleration (LWFA), emphasizing the need for imaging methods that are highly sensitive to refractive index gradients and can resolve spatial details on the micrometer scale and temporal events on the femtosecond scale.

1.2.3 Attosecond pulse generation

Attosecond pulse generation has recently garnered significant attention, highlighted by its recognition as the subject of the 2023 Nobel Prize in Physics [72]. The production of such ultra-short pulses relies heavily on the bandwidth-duration relationship, which states that shorter pulses require broader spectral bandwidths. This requirement makes High-order Harmonic Generation (HHG) a key mechanism for attosecond pulse generation, as it can produce light with the requisite bandwidth.

HHG is a collective molecular response to intense, coherent laser illumination. When a gas-phase ensemble of molecules is exposed to the strong and rapidly oscillating electric field of a high-intensity laser pulse, the binding potential that holds electrons near atomic nuclei becomes distorted. This distortion allows electrons to tunnel out of the potential barrier and be accelerated away by the laser's electric field. As the field reverses direction, the electrons are driven back toward their parent nuclei, where they may recombine. The energy accumulated during this trajectory is released upon recombination in the form of a high-energy photon, whose exact energy depends on the phase of the laser field at the time of electron tunneling. Due to the collective response of the ensemble and the periodicity of the process (occurring twice per cycle of the driving laser field), only the odd harmonics of the laser frequency constructively interfere, while others cancel out [73]. The resulting harmonics extend into the extreme ultraviolet (XUV) regime, enabling the generation of attosecond pulses.

The unique duration of attosecond (10^{-18} s) pulses opens up unprecedented opportunities for studying electron dynamics on their natural timescales [74]. For example, attosecond pulses have been used to measure the temporal delay in the photoelectric effect [75], providing insight into the internal dynamics of atoms and advancing our understanding of quantum processes.

While methods exist to characterize both the driving laser pulse [76, 77] and the resulting attosecond XUV pulse [78], no method currently allows for monitoring the generation process itself. Observing how the driving laser field ionizes the gas and is subsequently modified by the plasma it generates could significantly enhance our understanding of the process and enable further optimization of models of the generation process. Achieving such observations would require ultrafast single-shot videography with high spatio-temporal resolution and sensitivity to optical density gradients. Developing these capabilities would aid the further development of attosecond science, providing new insights into the underlying physics of ultrafast light-matter interactions.

The three examples given above underscore the potential of ultrafast single-shot videography to deepen our understanding of complex and dynamic processes. Each of these phenomena are at the intersection of cutting-edge science and technology, where capturing transient events with both high spatial and temporal resolution can lead to new discoveries. Beyond these cases, the need for ultrafast imaging extends into other fields, such as ultrafast chemical reactions [79, 80] and material science [81, 47], where imaging shock wave propagation and rapid phase transitions has the potential to lead to breakthroughs in understanding material properties under extreme conditions. Ultrafast imaging is therefore not merely a technical challenge but also a gateway to a deeper understanding of the physics that govern our universe.

2 The general solution

This section provides an overview of the general principles behind achieving ultrafast single-shot imaging, focusing on the use of ultrashort laser pulses and optical multiplexing. Before delving into the specific methods, I define and discuss the key metrics used to evaluate the performance of ultrafast imaging systems, such as spatial and temporal resolution, sensitivity, and dynamic range. Additionally, I will introduce imaging modalities relevant for studying transparent targets, including shadowgraphy, schlieren imaging, and interferometry, to provide a fundamental understanding of the techniques employed in ultrafast imaging.

2.1 The metrics

When comparing the utility of different systems, it is convenient to use numerical metrics that directly correlate with performance. The most relevant metrics depend on the system's intended purpose. In this section, the key metrics for ultrafast imaging are introduced.

2.1.1 Spatial resolution

Spatial resolution is often misunderstood as the number of pixels n_x [2] on an imaging sensor. While the pixel count is relevant and can sometimes be the limiting factor, it does not inherently define the spatial resolution of the imaging system as a whole.

The actual meaning of spatial resolution, r_x, is the minimum distance between two resolvable points. It can be interpreted as the inverse of the highest resolvable spatial frequency, often expressed as the number of line pairs per millimeter (lp/mm). Lord Rayleigh established an elegant criterion for resolution based on the diffraction patterns of coherent point sources, which produce concentric Airy rings when imaged (see Fig. 2.2a) [82, 83]. According to this criterion, two points are resolved when the central maximum of one diffraction pattern coincides with the first minimum of the other. [3] To generalize this definition for other signals, we calculate the corresponding critical Michelson contrast:

$$C_\mathrm{crit} = \frac{I_\mathrm{max} - I_\mathrm{min}}{I_\mathrm{max} + I_\mathrm{min}} = 0.1525, \tag{2.1}$$

where I_max and I_min are the intensities at the peaks and the local minimum between them, respectively. For Gaussian signals, this criterion translates to points being resolvable when separated by approximately 1.19 times their full width at half maximum (FWHM):

$$r_\mathrm{x} = 1.19 \cdot \mathrm{FWHM}. \tag{2.2}$$

The spatial resolution of an optical system can thus be determined either by identifying the closest features that satisfy the contrast criterion in Eq. 2.1 or by simply measuring the FWHM of the smallest observable feature and applying Eq. 2.2.

The limiting factor for an imaging system's *measurable* resolution can originate from the digital sensor, the optical system, or the physical target. This is illustrated in Fig. 2.2b, where three synthetic resolution measurement images are shown. In each case, a different parameter limits the measurable resolution, denoted as r_i.

The first constraint arises from the digital sensor. Its resolution is determined by the projected pixel size Δx, which is defined as the physical size of a reference object divided by the number of pixels it spans in a sensor image. Note that Δx is not the physical pixel size on the sensor but rather its effective size in the image plane. To resolve an array of spots, at least one pixel must separate adjacent spots, and each spot must occupy at least one pixel. This sets a lower bound for the resolution:

---

[2] For simplicity, only the x direction is discussed here, but the same principles apply to the y direction.

[3] Rayleigh's criterion has been shown to be slightly conservative compared to human perception of resolved spots [76], however, its elegant definition makes it a strong candidate for a universal resolution standard.

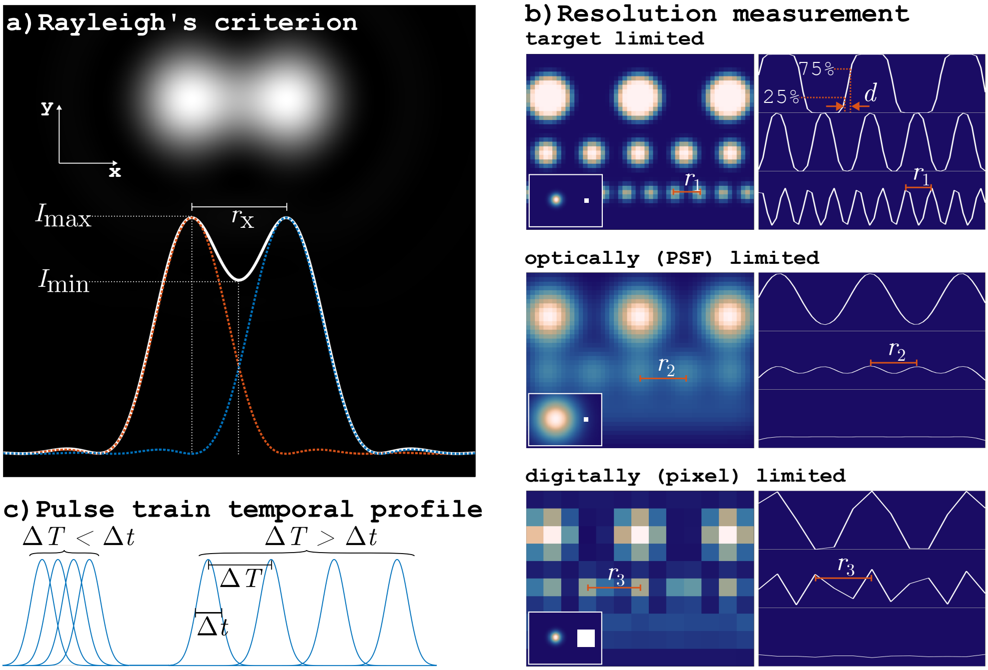


Figure 2.2: Illustrations of spatial and temporal resolution. (a) Two Airy patterns resolved according to Lord Rayleigh's criterion. The colored dotted lines represent the intensity profiles of the individual patterns, while the white solid line shows their summed intensity profile. The rings are only visible in the plots. (b) Three synthetically produced resolution measurements, each consisting of rows of spots. In each case, the measured resolution r_i is limited by a different factor. Within a row, the distance between adjacent spots equals twice the spot diameter. The images were generated by digitizing a convolution of a common ground truth image (identical in all three cases) of perfectly circular top-hat spots with a Gaussian point spread function (PSF). The resolution is determined as the minimum distance between resolvable spots. For each case, the size of the PSF and pixel is shown as an inset. When the size of the spots (the target) is the limiting factor, the resolution can be estimated by measuring the distance d over which the edge transitions from 75% to 25% of the maximum spot intensity and setting $r \approx 2.1d$. (c) Temporal intensity profiles of two pulse trains: one with strongly overlapping pulses ($\Delta T < \Delta t$) and another with non-overlapping pulses ($\Delta T > \Delta t$).

$$r_x \geq 2\Delta x.$$

The second constraint comes from the optical system and its point spread function (PSF), which describes how a point source is blurred when observed from the sensor plane. In the case of coherent, aberration-free illumination, the PSF takes the form of an Airy pattern. Applying Lord Rayleigh's criterion, the spatial resolution is given by:

$$r_x \geq \frac{0.51\lambda}{\mathrm{NA}},$$

where λ is the illumination wavelength and NA (numerical aperture) represents the light-gathering ability of the imaging system [84]. More generally, the resolution is limited by the FWHM of the PSF, leading to the criterion in Eq. 2.2.

Finally, the physical target itself may impose a resolution limit if it lacks sufficiently small or closely spaced features. When no small or closely spaced features are present in the image, the resolution can instead be estimated from the broadening of an edge. In this case, the 50% intensity drop-off distance $d_{50\%}$ (measured between 75% and 25% of the maximum edge intensity) can be used to approximate the resolution as:

$$r_\mathrm{x} \approx 2.1 d_{50\%}.$$

For a derivation of this expression, see Appendix A.

For systems designed to image transparent targets, traditional methods of resolution testing, such as using highly absorbing test patterns, may not be appropriate. Instead, resolution should ideally be measured in the same imaging modality and under conditions that closely resemble the intended application. This ensures that the resolution assessment reflects the actual performance of the system in its intended use case, avoiding overestimation or misrepresentation of its capabilities.

While spatial resolution is critical for assessing an imaging system's performance, it must be considered alongside the field of view (FOV). A larger FOV allows the system to capture more of the scene, enabling a broader perspective. In practice, there is often a trade-off between resolution and FOV, with higher spatial resolution achieved at the cost of a smaller FOV.

2.1.2 Temporal resolution

Temporal resolution is the minimum temporal separation between two events that can be observed as distinct. It should ideally be treated similarly to spatial resolution, meaning that it should be determined from data using Rayleigh's criterion. Such a rigorous approach is rare in the ultrafast imaging literature, though [85] provides a notable example. Unfortunately, temporal resolution is often conflated with the frame interval ΔT, which is the inverse of the frame rate. While this might be a reasonable definition for silicon-based sensors that acquire frames sequentially, it does not apply to more complex imaging systems and does not align with Lord Rayleigh's criterion.

In traditional high-speed cameras, the frame interval ΔT is strictly longer than the integration time Δt. However, in ultrafast imaging systems, especially those that rely on illumination by a train of short light pulses, ΔT can be shorter than the integration time Δt, as illustrated in Fig. 2.2c. Here, Δt represents the time during which light is collected to form a frame. For conventional sensors, Δt is the time the sensor is exposed to light, while for illumination-based methods, Δt corresponds to the duration of the illuminating pulses. [4]

---

[4] How Δt is defined depends on the temporal pulse shape, but assuming that the pulses are Gaussian and

By temporally overlapping these pulses, i.e., by setting $\Delta T < \Delta t$ (see Fig. 2.2b), the frame rate can be increased at the expense of the uniqueness of the data in each frame.

For ordinary cameras, this situation does not have a direct analog in the spatial domain because pixel size and spacing are inherently linked—pixels do not overlap. In the temporal domain, however, if $\Delta T < \Delta t$ data points do overlap, and the collected data can often be down-sampled without loss of information, effectively reducing the usable frame rate. To address this complexity, I propose that when temporal resolution cannot be rigorously measured, it should at least be estimated analogously to the spatial case as

$$r_\mathrm{t} \geq 2\Delta T,$$

i.e., twice the temporal spacing of illuminating pulses. This estimation strictly requires that $\Delta T \leq \Delta t$ - meaning *illuminating pulses do not overlap* - and aligns with the principle that at least two samples are required to resolve a feature [86].

Using this definition of temporal resolution, the two key metrics for the temporal domain become the temporal resolution r_t, analogous to spatial resolution, and the total integration time T, analogous to the field of view (FOV) in the spatial domain. Commonly, frame rate ($1/\Delta T$) and sequence depth (n_t) are used instead but they only reveal the total integration time ($T = (n_\mathrm{t} - 1)\Delta T$) and not the temporal resolution. While these metrics are indicative, they provide an incomplete picture, as they reveal little about how fast dynamics the system's is actually able to capture accurately. For example, if a system employs many long, overlapping illuminating pulses, it may achieve a high frame rate and an impressive sequence depth, but the resulting frames will carry nearly identical information and the temporal resolution will be poor due to the long pulse durations. 100 pulses with a duration of $\Delta t = 1$ ps could for example be separated by just $\Delta T = 10$ fs, to obtain a staggering 100 Tfps video sequence over 100 frames. Adjacent frames would however have an information overlap around 99% and the total integration time would only be ~1 ps. In contrast, temporal resolution r_t and total integration time T directly reflect how well and for how long a transient event can be recorded—qualities that ultimately determine the system's utility in capturing dynamic phenomena.

Another crucial aspect to keep in mind when imaging trasient events is the characteristic velocity v of the moving object, as it directly links temporal and spatial resolution. A key objective for scientific videographic methods is to achieve motion blur-free imaging, as emphasized in e.g. [87, 88, 89]. To accomplish this, the object must not move more than one pixel, [5] Δx during the integration time:

---

requiring that the illumination pulses are resolvable (according to Lord Rayleigh's criterion) when $\Delta t = \Delta T$ (back-to-back pulses) results in $\Delta t = 1.19\mathrm{FWHM}$, where FWHM is the pulse's full width at half maximum.

[5] Allowing a movement of up to one pixel ensures that a point is never imaged by more than two pixels (per

$$v\Delta t \leq \Delta x. \tag{2.3}$$

Less intuitive but entirely analogous to motion blur is the issue of ”frozen motion”, as illustrated in Fig. 2.3. Points moving at different speeds are captured individually in Fig. 2.3a, while in Fig. 2.3b, the response to all three points is visualized, highlighting the symmetry between motion blur and frozen motion. The yellow point moves too fast for the camera, causing motion blur, while the blue point moves too slowly, appearing stationary (frozen motion).

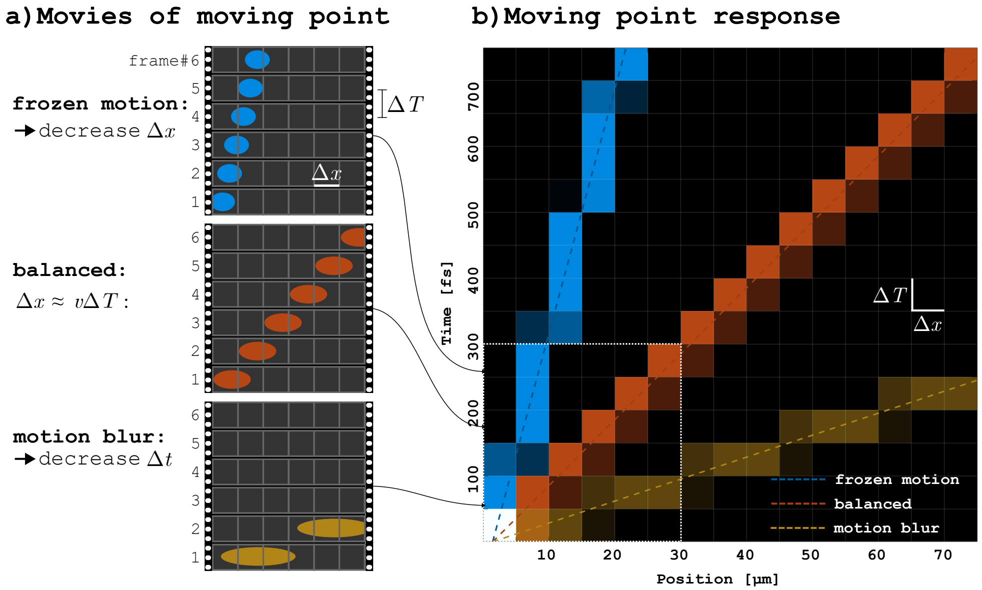


Figure 2.3: Recording a moving point. (a) Six single-row frames of three moving point sources (blue, orange, and yellow) with different velocities in the positive x-direction. The camera sensor is one-dimensional, capturing one row per frame. The integration time Δt introduces motion blur, while the inter-frame time ΔT determines whether motion is frozen. These sequences correspond to the white box in (b). (b) Simulated camera response to the same three point sources, illustrating the symmetry between spatial and temporal resolution. The x-axis represents position, while the y-axis represents time. Pixel size is $\Delta x = 5 \text{ m}$, and the inter-frame time is $\Delta T = 50 \text{ fs}$. The color intensity reflects how long a point remains within a pixel. The yellow source ($v = c$) moves too fast and appears blurred (motion blur), the blue source moves too slowly and appears static (frozen motion), while the orange source is optimally sampled, showing neither effect.

A general audience often reacts negatively to cases like the yellow point, demanding shorter integration times to reduce motion blur. In contrast, scenarios like the blue point might be admired for their ability to ”freeze” motion. However, as Fig. 2.3 demonstrates, these situations are fundamentally equivalent. Both motion blur and frozen motion degrade the system’s ability to capture dynamics accurately: motion blur indicates insufficient tempo-

---

direction). Requiring that a moving point be captured by only one pixel at all times would be impractical, as even a slow-moving object could randomly transition between adjacent pixels during a short integration time.

ral resolution, while frozen motion arises due to inadequate spatial resolution to detect movement.

To avoid frozen motion, the object must move at least one pixel between frames, which requires:

$$v\Delta T \geq \Delta x. \tag{2.4}$$

Ultimately it is Equations 2.3 and 2.4 that dictate how pixel size Δx, integration time Δt and frame interval ΔT must be related to prevent both motion blur and frozen motion. However, for the ideal case where the system operates at its temporal resolution limit (i.e. when $\Delta T = \Delta t$), and the collected data is down-sampled to match the actual resolutions, the spatial and temporal resolutions are given by:

$$r_\mathrm{t} = \gamma \cdot 2\Delta t = \gamma \cdot 2\Delta T, \quad r_\mathrm{x} = \gamma \cdot 2\Delta x,$$

for some down-sampling factor γ. Since the Rayleigh resolution criterion that defines both r_x and r_t is somewhat conservative, γ should likely be slightly larger than one to avoid unwanted loss of information. The exact value of γ is not critical, only that the temporal and spatial dimensions should be treated equally. Substituting these relationships into Equations 2.3 and 2.4, which state the criteria for motion blur and frozen motion, respectively, yields the following relationship for a system with balanced spatiotemporal resolution:

$$r_\mathrm{x} = v \cdot r_\mathrm{t}. \tag{2.5}$$

In practice, perfectly matching spatial and temporal resolution is very challenging. However, avoiding significant mismatches is essential. If r_x is much smaller than $v \cdot r_\mathrm{t}$, motion blur will degrade the image, indicating that the temporal resolution must be enhanced (i.e. r_t made smaller). Conversely, if $v \cdot r_\mathrm{t}$ is much smaller than r_x, motion will be undetectable, indicating oversampling in time, indicating that the spatial resolution must be enhanced. To ensure accurate capture of dynamic phenomena, spatiotemporal resolution should be as closely aligned as possible.

For systems intended to image laser-induced phenomena (i.e., $v = c = 3 \cdot 10^8$ m/s) as discussed in Sec. 1.2, such as laser filamentation, laser wake-field acceleration, and attosecond generation, with a suggested temporal resolution of approximately 20 fs, Eq. 2.5 implies that a spatial resolution of approximately 6 µm is desirable.

2.1.3 Sensitivity

Spatial and temporal resolution, as well as field of view and total integration time, are crucial parameters for determining the capability of an imaging system, but without sufficient sensitivity to a signal from the target, they become meaningless. In ultrafast imaging, where interframe times are on the order of a few picoseconds (1 ps = 10^{-12} s) or less, target emission processes are generally too slow and weak to be relevant. As a result, the signal must typically be imprinted on and mediated by external illumination—except in cases where the illumination itself is studied object, as in [90, 91]. For transparent targets, such as those listed in Sec. 1.2, the three most important imaging modalities are shadowgraphy, schlieren imaging, and interferometric imaging.

In shadowgraphy, the shadow cast by the target is observed. Shadows can result from absorption, as with opaque objects, or from the lensing effect of transparent targets. For opaque objects shadowgraphy is highly sensitive, since it produces a strong and clear signal, while for transparent targets the signal will be much weaker. A transparent target induces angular deflections in incident light rays according to the optical density distribution of the target, as governed by Snell's law. If the deflection is homogeneous, i.e., if the derivative of the deflection is zero, shadowgraphy will detect no signal. However, if the deflection is inhomogeneous, certain sensor areas will receive more light while others receive less, resulting in a shadow pattern that serves as the signal from the target. [6] The more inhomogeneous, i.e. the greater the derivative of the deflection is, the stronger the signal will be.

Schlieren imaging detects even homogeneous deflections by selectively discarding light based on the magnitude and direction of the deflection. Traditionally, this is achieved using a spatial filter at the focus of an imaging lens, where the signal from deflected light either increases or decreases in intensity depending on the direction of deflection and the geometry of the filter [92]. A modern variation, background-oriented schlieren (BOS), quantifies perceived distortions in the target's background rather than directly filtering light [93, 94]. Dark-field microscopy can also be considered a form of schlieren imaging, as it relies on light deflection to generate contrast [95]. Since schlieren imaging directly measures light deflection rather than its second derivative, it inherently offers higher sensitivity than shadowgraphy.

In Fig. 2.4, BOS has been implemented in a home-office setup using a webcam as a sensor, a sinusoidal pattern displayed on a computer monitor as the background, and the rising hot air from a tealight as the target. [7] The signal manifests as slight perturbations in the striped pattern, which can be readily analyzed through Fourier methods (see Sec. 3.2).

---

[6] For example, sunlight illuminating a seabed through wavy water produces a shadowgraphic pattern of lighter and darker areas.

[7] Every parent should try this experiment, as it has been proved by example that children as young as five take great delight in observing the invisible.

While observing the target in a shadowgraphic configuration yields no visible signal, the schlieren setup reveals a clear and strong signal, exposing the turbulent rising air. The stark contrast between the two signal images in the figure highlights the significantly greater sensitivity of schlieren techniques compared to shadowgraphy when imaging transparent targets.

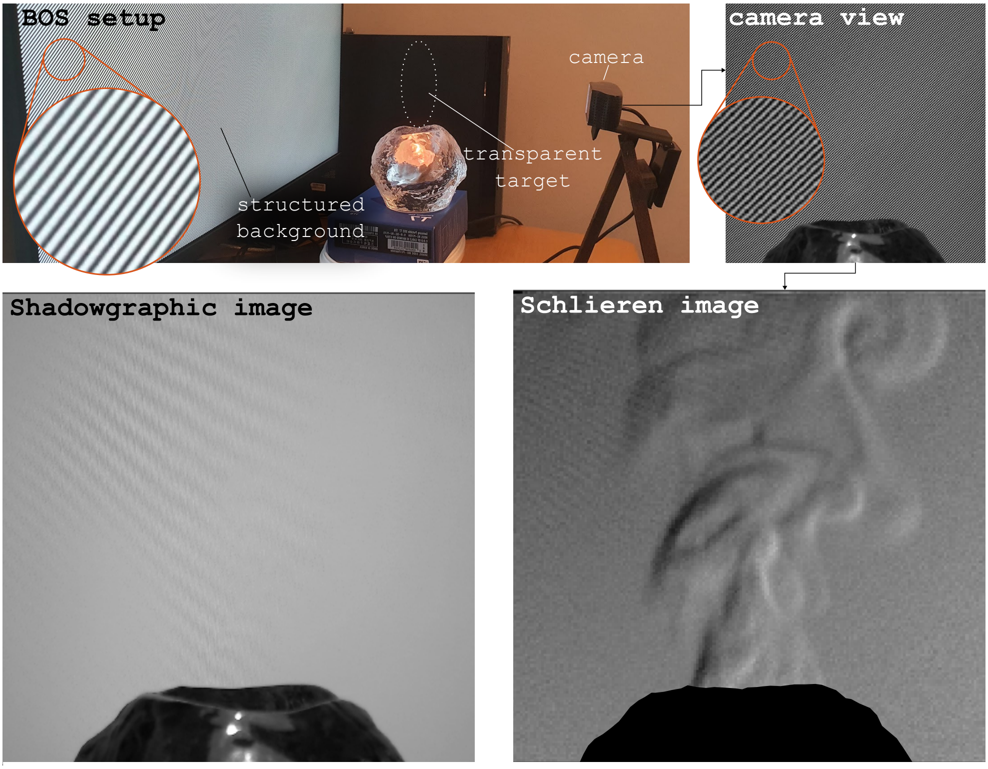


Figure 2.4: A simple implementation of background-oriented schlieren using a structured background. The schlieren image was reconstructed from the camera view image, while the shadowgraphic image was obtained using a uniform background. The comparison highlights the increased sensitivity of schlieren imaging for detecting transparent phenomena.

The most sensitive modality, however, is interferometric or holographic imaging. These methods detect not only light deflection but also the optical delay induced by the target. In interferometry, a probe pulse that has passed through the target is made to interfere with a reference pulse. [8] By comparing the resulting interference patterns with and without the target, the optical delay (i.e., the added optical path length) introduced by the target can be detected and quantified as a phase shift. These methods are extraordinarily sensitive, capable of detecting variations in path length as small as 1 nm [96]. Consequently, interferometric methods represent the gold standard for imaging transparent objects.

---

[8] In interferometry, the reference pulse is typically split off from the probe pulse before target interaction, whereas in holography, the reference pulse is usually an entirely separate pulse.

2.2 Short laser pulses

With electronics being too slow to detect the ultrafast phenomena we aim to observe, the solution is to shift from relying on electronic speed to relying on the speed of light. Instead of increasing temporal resolution by making detectors faster, we can instead shorten the duration of the incoming light. By doing so, the time that the detector receives light is reduced, even if the detector itself is slow. Since off-the-shelf lasers can generate pulses as short as a few femtoseconds (1 fs = 10^{-15} s), it becomes possible to achieve the extremely short integration times required for imaging ultrafast events.

In principle, this is achieved by sending a short probe pulse through the event of interest and then recording it with a camera sensor. If the event affects the probe pulse, this interaction is captured as a signal. By iteratively repeating the process with a slightly increased delay between the event onset and the probe pulse, a video of the event can be reconstructed. Such pump-probe experiments have been used, for example, to capture image sequences of light in flight [97]. However, pump-probe experiments require the event to be repeatable, which is not the case for stochastic or destructive processes.

For these types of events, the entire process must be recorded in a single shot to capture its evolution accurately. This is accomplished by illuminating the event not with a single pulse, but with a train of pulses, each probing the event at distinct times. In this way, each pulse carries information corresponding to a different phase of the process. The challenge then becomes ensuring that the optically distinct information carried by each pulse remains distinguishable even after they are captured by the camera. Achieving this requires overcoming the limitations of traditional systems through optical multiplexing.

2.3 Optical multiplexing

Multiplexing is the process of simultaneously sending more than one signal through a single physical channel. For the process to be meaningful it must be accompanied by de-multiplexing, the inverse process of separating multiplexed signals. De-multiplexing takes place on the other side of the channel and either allows a single receiver to read multiple signals, or multiple receivers to each read their designated signal. If multiplexing is not followed by de-multiplexing, the only thing it achieves is the mixing-up of signals and subsequent loss of information. Hence, the term multiplexing often implicitly also includes de-multiplexing, and that is generally how I will use the term.

In principle it would not be wrong to call the acts of hearing and decoding two simultaneous voices multiplexing and de-multiplexing, respectively. [9] However, the term is in

---

[9] The situation might arise when you overhear a more interesting conversation at the table next to yours, while still managing to somewhat follow what your table mate is saying.

general reserved for more technical applications, specifically within telecommunication. In the 1870's the predecessor of all telecommunication, telegraphy, was an established technology, and the demand for telegrams was growing high enough that the wires became a limiting factor for telegraphy companies. Each telegram consisted of a sequence of beeps (Morse code), which could not be interrupted, without the message being lost. Hence, only one message could be transmitted at the time. Adding more wires to meet the increasing demand was costly and it was in this context that Elisha Gray in 1875 invented a multiplexer, enabling dual signal transmission by sending beeps of different pitches. These signals could then be separated by tone at the receiving end, allowing multiple messages to be sent simultaneously without loss of information [98].

It turns out that the challenge with ultrafast imaging is similar to the challenge with telegraphy in the 1870's. In ultrafast imaging it is not wires that are the scarce resource that has to be shared by multiple signals, but camera sensors. Recording a single ultrashort optical signal with a camera sensor is straightforward: the sensor receives the optical signal, converts it to electrons in its pixels, and reads out the voltage of each pixel sequentially. Sensor speed is irrelevant in this scenario, as it has all the time it needs. However, recording just two light pulses with temporal separations shorter than the sensor's readout time immediately becomes challenging. Just as telegraphy companies could not meet demand the sensor cannot meet the brief, but ultrahigh, demand for optical storage.

To overcome this limitation, some optical trick has to be applied to the light pulses before they reach the camera. These tricks ensure that distinct optical signals remain distinguishable even after being recorded by the sensor. In the field of ultrafast imaging, optical multiplexing techniques can be broadly categorized into: (1) spatial division multiplexing, (2) wavelength division multiplexing, (3) compressed sensing-based multiplexing, and (4) spatial frequency division multiplexing. The technique used in this work, FRAME, falls into the last category.

3 Specific solutions

This section provides an overview of several ultrafast imaging techniques, focusing on their fundamental principles of operation, key achievements in resolution, frame rate, and imaging modality, as well as their respective strengths and limitations. For those exploring the broader field of ultrafast imaging, this overview offers a comprehensive entry point. General readers, however, may still gain a solid introduction to the field by focusing on the subsection introductions and examining the figures.

3.1 Spatial division multiplexing

The common principle underlying all Spatial Division Multiplexing (SDM) techniques is that temporally distinct events are recorded at spatially distinct regions of the detector. In practice, probe pulses are sent through the target volume and then directed to unique locations on the detector. The primary challenge lies in spatially separating the probe pulses. To achieve this, the pulses must differ in some parameter—such as wavelength, polarization, or another distinguishing characteristic. These differences are then exploited to sort the pulses based on the chosen parameter, allowing them to be mapped to specific regions of the detector.

3.1.1 Sequentially timed all-optical mapping photography

Sequentially Timed All-optical Mapping Photography (STAMP) was first introduced in 2014 by Nakagawa et al. [99]. This technique marked the advent of ultrafast single-shot imaging and relied on spatial division multiplexing. [10] The parameter used to sort the probe pulses was wavelength, and the basic principle can be summarized in four steps, illustrated and described in Fig. 2.5.

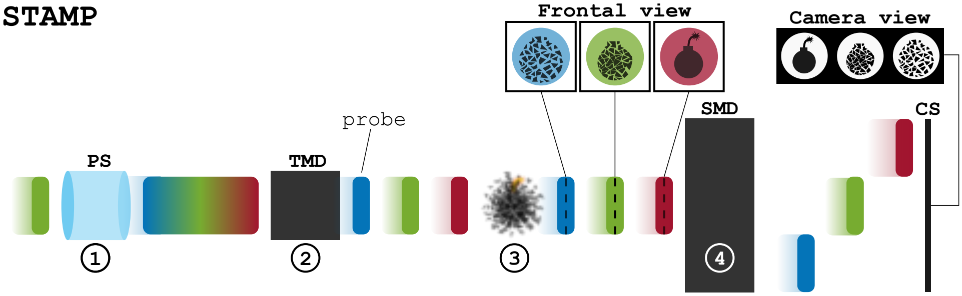


Figure 2.5: The principle operation of STAMP. (1) An initially short, broadband pulse is stretched in time to match the desired total integration time. (2) The pulse is divided into a train of short, spectrally unique probe pulses. (3) Each probe pulse interacts with the target at a unique time, imprinting information corresponding to the target's state at the time of interaction. (4) The probe pulses are spatially directed to distinct detector locations based on their wavelengths. PS: pulse stretcher; TMD: temporal mapping device; SMD: spatial mapping device; CS: camera sensor.

Since its introduction, STAMP has been demonstrated in numerous studies, with continual refinements to its technical implementation. In its original version, the system used a Temporal Mapping Fevice (TMD) to divide the initial stretched pulse into temporally distinct probe pulses. The TMD consisted of two gratings, two lenses, two wave plates, and a spatial light modulator. A Spatial Mapping Device (SMD) was then employed to direct the probe pulses to distinct detector locations. The SMD included another grating,

---

[10] While it was not the very first single-shot ultrafast imaging technique (see [100]), it marked the beginning of a rapid surge in the development of such techniques.

a cylindrical mirror, and a custom-built periscope array with six channels. Each channel redirected a probe pulse vertically and laterally, eventually guiding it to the detector. This configuration enabled the acquisition of six shadowgraphic images at an unprecedented frame rate. However, the group later acknowledged challenges with scalability, noting, "it is not easy to manufacture a periscope that can manage much larger spectral components," and concluded that "another approach is needed" [101].

This alternative approach, presented in the same paper from 2015, replaced both the TMD and SMD (and their numerous components) with a simpler configuration. A Diffractive Optical Element (DOE) was used to spatially split the initial pulse into sub-pulses, while an angled bandpass filter selected distinct spectral—and therefore temporal—components of each sub-pulse. In this configuration, named Spectral Filtering (SF-) STAMP, the pump pulse interacted with the target before being divided into sub-pulses. However, this simpler setup introduced a new issue: the sub-pulses were no longer equidistant in time. Furthermore, intensity variations within the illumination, a long-standing problem across all STAMP versions, persisted.

Both of these issues were resolved in 2020 by another research group, which incorporated an Acousto-Optic Programmable Dispersive Filter (AOPDF-STAMP). This innovation allowed precise temporal control and spectral selection, ensuring equidistant sub-pulses with uniform intensity [102]. The same study also demonstrated STAMP with in-line holography, making the technique phase-sensitive for the first time.

In 2024, a third research group achieved optical phase-sensitive imaging by integrating STAMP with diffraction phase microscopy [103]. In this configuration, a reference pulse was split from each probe pulse after interaction with the target using a grating. The reference pulse was spectrally filtered in the optical Fourier plane to remove high-frequency components and then overlapped with its corresponding probe pulse on the detector, generating an interference pattern that encoded phase information about the target.

The highest reported frame rate for STAMP systems is 7.52 Tfps, corresponding to an interframe time of 133 fs [104]. However, the probe pulses had a duration of 464 fs, resulting in 331 fs temporal overlap and an effective frame rate of approximately 2.2 Tfps when considering the integration time. This study also reported the highest number of frames achieved in a STAMP system, namely 25. Spatial resolution, though not directly quantified in any STAMP paper, can sometimes be estimated from the smallest structures visible in the data. In the original STAMP paper, the spatial resolution was approximately 3 µm (330 lp/mm) over a field of view (FOV) of 93×93 µm^2 [99]. In subsequent studies, it was approximately 8 µm (125 lp/mm) over $207 \times 135 \times 93$ µm^2 for SF-STAMP [101] and 15 µm (67 lp/mm) over $190 \times 245 \times 93$ µm^2 for AOPDF-STAMP [102].

STAMP systems have been applied to a range of phenomena, including laser ablation [99, 105, 102], lattice vibrations [99], air breakdown [101, 102], laser-induced plasma filaments

[101, 103], ultrafast phase transitions [81], laser-induced shock waves [106], and the opening and closing of an optical Kerr gate [102].

An inherent characteristic of STAMP is the need for wavelength differences between the probe pulses, typically on the order of tens of nanometers. While this is unlikely to impact many applications, it introduces a theoretical risk that variations in the observed signal between frames could stem from differences in the probing pulses rather than changes in the target. This issue is particularly relevant in cases where the spectral sensitivity of the target could influence the interpretation of results, underscoring the importance of carefully considering potential wavelength-dependent effects in experimental designs. Despite this, STAMP is clearly a powerful tool for capturing ultrafast dynamics with single-shot acquisition.

3.1.2 Framing imaging based on non-collinear optical parametric amplification

Framing Imaging based on Non-Collinear Optical Parametric Amplification (FINCOPA) was introduced by Zeng et al. in two papers in 2020 [88, 107]. FINCOPA is an SDM technique that achieves single-shot ultrafast imaging by sequentially mapping temporally distinct parts of a probe pulse to unique detectors using Non-Collinear Optical Parametric Amplification (NCOPA) crystals. The principle of operation is schematically presented in Fig. 2.6.

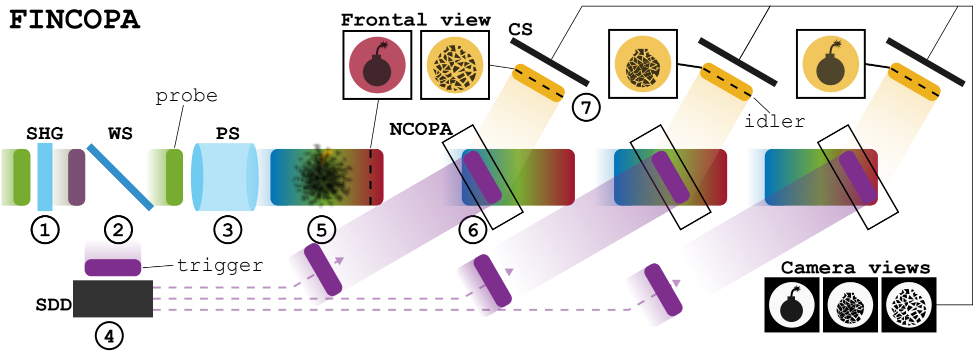


Figure 2.6: The principle operation of FINCOPA. (1) An initial short pulse passes a second harmonic generator (SHG) to form a probe pulse and a frequency doubled trigger pulse. (2) The two pulses are separated. (3) The probe pulse is stretched in time to match the desired total integration time. (4) The initial trigger pulse is further split into multiple trigger pulses, which are temporally controlled to match the desired integration times. (5) The probe pulse interacts with a target, continuously imprinting the dynamics of the scene. (6) The probe pulse is sequentially overlapped with the triggering pulses in distinct NCOPA crystals, generating idler pulses that correspond to temporally distinct parts of the probe pulse. (7) Each idler pulse is mapped to its own detector. SHG: second harmonics generator; WS: wavelength separator; SDD: split and delay device; PS: pulse stretcher; NCOPA: noncollinear optical parametric amplifier; CS: camera sensor.

FINCOPA has been demonstrated with a frame rate of up to 15 Tfps ($\Delta T = 67$ fs) over four shadowgraphic frames, using 50 fs triggering pulses. The spatial resolution was measured at 12 µm (83 lp/mm) by imaging the propagation of a plasma grating, with a field of view of approximately $900 \times 600\,\mu\mathrm{m}^2$.

In terms of studied phenomena, FINCOPA has been applied to laser-induced plasma gratings and Ultrafast Intensity-Rotating Optical Fields (UIROF), which rotate at rates of up to 1.9 trillion revolutions per second. The UIROF sequences are particularly noteworthy, as they provide a clear visualization of how a circular light field with a low-intensity central streak rotates distinctly from frame to frame.

One of FINCOPA's key strengths lies in its use of four separate detectors, which eliminates the trade-off between sequence depth and spatial resolution. This enables the resolution of small structures without compromising the number of frames captured. However, this approach comes with certain drawbacks. Each additional frame requires an additional detector, which increases both the cost and the complexity of the setup. Furthermore, the later frames in a sequence are encoded in parts of the probe pulse that have passed through more optical components—such as lenses and NCOPA crystals—than the earlier frames. This effectively stretches the probe pulse in time for each additional frame, while the triggering pulses remain unaffected, leading to variations in readout conditions across the sequence. Despite these challenges, FINCOPA has demonstrated impressive capabilities, achieving high spatial and temporal resolution.

3.1.3 Framing integration ultrafast photography

Framing Integration Ultrafast Photography (FIP), also referred to as Frequency domain Integration Sequential Imaging (FISI), is a spatial division multiplexing technique introduced by Zhu et al. in 2021 and further developed in 2022 [108, 109]. It achieves ultrafast imaging by encoding temporal information through spatial delays introduced in the probe pulse, enabling simultaneous acquisition of multiple shadowgraphic frames in a single shot. The FIP system integrates the temporal delay unit and imaging structure into a simple and compact setup. Its operational principle is illustrated in Fig. 2.7.

In the referenced studies, FIP demonstrated a maximum frame rate of 12.6×10^{12} fps ($\Delta T = 79$ fs) over six frames, with a spatial resolution of approximately 50.8 lp/mm or 20 µm. The reconstructed frames were 2.6 mm × 3.7 mm in size, and the effective signal-detection area was limited to 1.0 mm × 0.35 mm. When measuring the spatial resolution, a stationary target was used.

The authors also reported an intrinsic spatial resolution of 108 lp/mm, a value dependent on the system's optical magnification ($M = 0.41\times$). However, intrinsic resolution does not represent the actual resolvable detail in recorded images and is not a standard performance metric for imaging systems. The observable (extrinsic) spatial resolution of

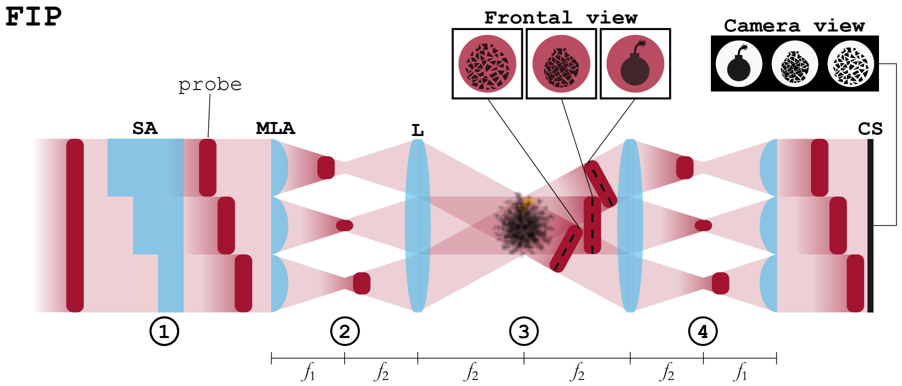


Figure 2.7: The principle operation of FIP. (1) Spatially distinct regions of a short initial pulse are individually delayed by different temporal offsets using a step array made of glass. (2) The delayed pulses are collimated and directed towards the target using a microlens array and an ordinary lens. (3) The probe pulses interact with the transient target, encoding the dynamics of the scene. (4) The probe pulses are collimated and mapped to distinct detector regions by a second lens and a matching microlens array. SA: step array; MLA: microlens array with focal length f_1; L: lens with focal length f_1; CS: camera sensor.

50.8 lp/mm reflects the system's practical capability in its operational configuration and is the relevant measure for assessing FIP's imaging capabilities. In both studies where FIP was demonstrated, the technique was used to image laser-induced plasma channels. This remains the only phenomenon studied using this method, underscoring its early stage of application in ultrafast imaging research.

FIP has certain limitations that affect its performance, particularly when imaging transparent targets, such as plasma filaments. The technique exhibits low sensitivity to refractive index changes in such cases, necessitating background division to make the signal visible. This low sensitivity is due to the shadowgraphic modality and results in reduced contrast, which can degrade the observable spatial resolution compared to the values measured using highly absorbing resolution targets. An underlying issue is that the spatial resolution is not measured in the same modality as the intended targets, complicating the evaluation of its performance under realistic experimental conditions. An issue directly related to the geometry of the setup is that temporally distinct pulses will probe the target from slightly different angles, yielding a somewhat stereoscopic view that makes direct comparisons between frames difficult. Nevertheless, the technique's reliance on commercial optical components and the streamlined integration of the delay unit in the imaging system suggest potential for further optimization.

3.2 Wavelength division multiplexing

Wavelength division multiplexing (WDM) operates on the principle that temporally distinct events are probed by spectrally distinct light in such a way that they can be separated based on the probing wavelength. Unlike spatial division multiplexing, where pulses are fully spatially separated on the detector (as in the wavelength-dependent spatial division multiplexing technique STAMP), WDM allows the pulses to overlap spatially on the detector. This overlap requires that either the detector itself or the pulses be locally organized—at the pixel level—in a manner that enables the spectral separation of signals.

3.2.1 All-optical photography with raster

All-Optical Photography with Raster (OPR) is a WDM technique introduced in 2021 by Zhu et al. [89] in 2021. OPR maps the temporal dimension to the spectral dimension, which is subsequently mapped to specific positions on the detector. This approach takes advantage of the Nyquist sampling theorem, which states that a signal (in this case, a 2D image) can be reconstructed if each frequency component is sampled at least twice per period [86]. Consequently, the scene can be discretely sampled and reconstructed with a resolution determined by twice the sampling distance. The basic operational principle of OPR can be summarized in five steps, schematically illustrated in Fig. 2.8.

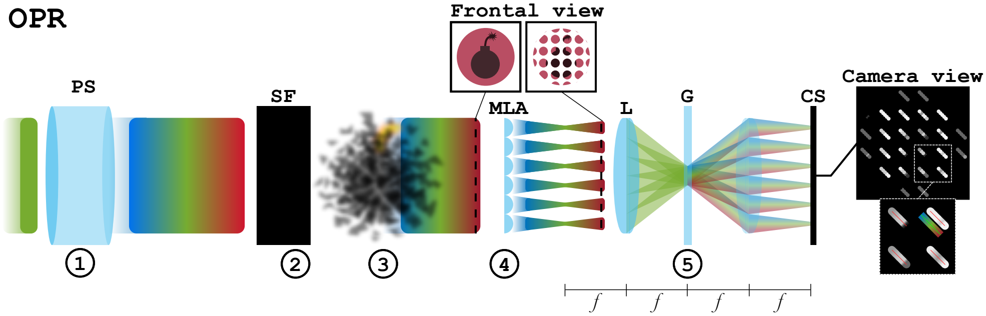


Figure 2.8: Operational principle of OPR. (1) An initially short, broadband probe pulse is temporally stretched. (2) The pulse is spectrally filtered to match the desired total integration time. (3) The pulse interacts with a transient target, continuously imprinting temporal information about the dynamic scene. (4) The pulse is spatially discretized in the x-y plane to form a raster of spectrally broad sub-pulses. (5) Each sub-pulse, comprising a wide range of wavelengths, is spatially dispersed according to wavelength and recorded by a camera sensor. PS: pulse stretcher; SF: spectral filter; MLA; microlens array; L: lens with focal length f, G: grating; CS: camera sensor.

The sensor image consists of an array of short lines, with each line corresponding to a position in the x-y plane and each position along the line encoding a specific wavelength—and thus a specific moment in time. Using a Fourier-based reconstruction algorithm, individual frames representing unique temporal instances can be extracted. This process relies on calibrating the spectral dispersion introduced by the grating and the setup geometry to establish a mapping between detector position, wavelength, and time. The calibration ensures that each reconstructed frame corresponds to a discrete temporal slice of the transient event, enabling the visualization of dynamic processes throughout the integration window.

In the referenced study the temporal mapping (steps 4 and 5 in Fig. 2.8) is implemented using a 300 × 300 microlens array and a grating positioned at the center of a 4f imaging system. The microlens array spatially discretizes the probe pulse, while the grating disperses the spectral components, mapping temporal instances to distinct detector locations. Although the article does not explicitly address how unwanted diffraction orders from the grating are filtered, it is reasonable to assume that this issue was effectively managed in the experimental setup. This configuration transforms the focal spot of each microlens into a temporally resolved line on the detector. By combining this spatial and spectral information, the system encodes both spatial and temporal details of the target interaction into the detector image. After reconstruction, the resulting images are of shadowgraphic modality, capturing intensity variations caused by the interaction between the probe pulse and the target.

The study reported a frame rate of 2 Tfps ($\Delta T = 500$ fs) over 12 frames, with a spatial resolution of 90 lp/mm (11 µm). The reconstructed images are reported to have dimensions of 1236 × 1626 pixels. However, given the composition of the microlens array, the number of effective sampling points for each image is 300 × 300, meaning that the images can at least be downsampled to this size without any loss of information. The field of view is not reported, but can be estimated to be 5 × 7 mm. Additionally, the spatial resolution was measured from a single (non-multiplexed) acquisition using a stationary target, while the temporal resolution is not specified. The probe pulse, centered at 800 nm with a bandwidth of 28 nm, interacted with laser-induced plasma strings in air and plasma dynamics in glass, which served as the study's targets.

As with other wavelength-dependent techniques, it is important to consider that different wavelengths may interact slightly differently with the target. Furthermore, while it is theoretically possible to calculate the temporal mapping of each sensor pixel (based on the grating parameters and setup geometry) and thereby estimate the frame interval, determining the degree of temporal overlap between adjacent frames is not trivial and is not addressed in the article.

3.2.2 Chirped spectral mapping ultrafast photography

Chirped Spectral Mapping Ultrafast Photography (CSMUP) is a straightforward ultrafast imaging technique introduced by Yao et al. in 2021 [110]. It operates by mapping temporal information encoded in a chirped broadband laser pulse into the spectral domain, which is then recorded using a hyperspectral camera. In this setup, each pixel consists of 25 subpixels, each covered by a unique band-pass filter (BPF) to selectively capture distinct spectral components. The basic principle is summarized in three steps in Fig. 2.9.

The resulting 2D sensor image contains both spatial and spectral information, from which shadowgraphic images of the temporal evolution of the target can be reconstructed frame by frame. The experimental steps are sketched in Fig. 2.9.

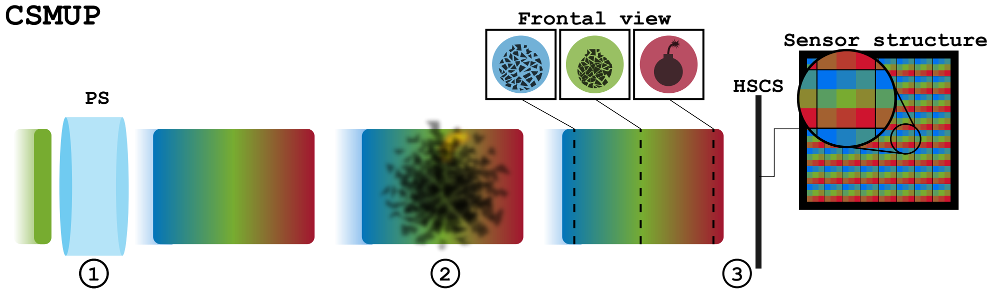


Figure 2.9: Operational principle of CSMUP. (1) An initial short probe pulse is temporally stretched to match the desired total integration time. (2) The probe pulse interacts with the target, imprinting temporal information corresponding to the ultrafast dynamics. (3) The pulse is captured by a hyperspectral camera, which encodes spectral – and hence temporal– components based on the sub-pixel band-pass filter arrangement. PS: pulse stretcher; HSCS: hyperspectral camera sensor.

In the two available studies, CSMUP produced video sequences of 25 frames at a frame rate of 250 Gfps ($\Delta T = 500$ ps). The reconstructed frames have a reported spatial resolution of 833 nm and pixel dimensions of 217 × 409 [110, 111]. Based on the presented images, the field of view can be estimated to be approximately 60×60 μm^2. The temporal resolution is influenced by the degree of overlap between frames, which is determined by the spectral overlap of the sub-pixel BPFs. While this overlap is not explicitly quantified in the articles, the presented spectral profiles suggest significant variation across the spectral range, ranging from low to substantial. This variation directly impacts the precision of the reconstructed temporal dynamics.

In the referenced studies, CSMUP was used to observe femtosecond laser ablation in materials like silicon and gold. While the captured data showed trends consistent with theoretical models of laser-material interaction, some discrepancies in quantitative agreement highlight the need for further refinement and validation. The technique faces limitations, including spectral-temporal crosstalk and unequal temporal intervals between frames due to the non-linear chirp of the probe pulse and varying spectral distance between sub-pixel BPF's. Despite these challenges, its simplicity, compact configuration, and reliance on commercially available components make CSMUP an appealing option for single-shot ultrafast imaging.

3.3 Compressed sensing-based multiplexing

Compressed sensing (CS) was introduced in 2006 by Candès et al. [112] and Donoho [113] as a method to reconstruct signals from far fewer samples than the Nyquist–Shannon theorem would suggest. Ordinarily, undersampling a signal would yield infinitely many possible solutions, making accurate reconstruction impossible. However, CS leverages prior knowledge about the signal to select a unique and stable solution.

In the case of natural images, the prior knowledge is that they are sparse in the Fourier domain. Rather than allowing any solution that fits the undersampled data, CS techniques iteratively approach the *sparsest* of all possible solutions through optimization-based algorithms. Requiring sparseness effectively narrows the solution space and allows the original image to be reconstructed accurately despite undersampling.

When applied to ultrafast imaging, compressed sensing allows multiple frames to be recorded within a single shot by intentionally undersampling and then numerically reconstructing the image sequence. Since the development of the first CS-based ultrafast imaging technique, numerous variations have appeared—often bearing distinct names. For simplicity, however, I will discuss all such variations in the same section.

3.3.1 Compressed ultrafast photography

Compressed Ultrafast Photography (CUP) was first demonstrated in 2014 by Gao et al. [90], just months after the first STAMP paper. CUP is described as a single camera *snapshot* technique rather than a single-*shot* imaging technique, as it does not require active illumination in principle. Unlike other ultrafast imaging methods, CUP does not map the temporal dimension of a dynamic scene to a specific property of the illumination. Instead, it uses a streak camera to encode temporal information. For this reason, one could argue that CUP deviates from the general solution to the ultrafast imaging problem outlined in Sections 2.2 and 2.3 of this chapter. However, in practice, CUP shares significant similarities with other ultrafast imaging techniques, as can be seen in in Fig 2.10, where its principle of operation is illustrated.

In CUP, the light from the dynamic scene that is ultimately recorded by the streak camera is conceptualized as a continuous 3D data cube, with spatial information encoded along the x- and y-axes and temporal information along the z-axis. As this data cube approaches the detector within the streak camera, it is sheared in the y-direction, establishing a correlation between the vertical detector position and time. However, this correlation is not one-to-one, as multiple temporal instances are recorded on the same detector row, resulting in undersampling. Specifically, subsequent frames share all but one row on the detector, with the frame rate determined by the ratio of the shearing velocity to the pixel size. This undersampling necessitates the use of compressive sensing-based reconstruction algorithms, whose primary goal is to recover the original data cube in a discretized form while maximizing spatio-temporal resolution and preserving the intricate details of the dynamic event under investigation. To achieve this, a random binary pattern is added to spatially encode the dynamic scene. This pattern introduces unique spatial signatures, enabling the separation of overlapping temporal and spatial information during the reconstruction process.

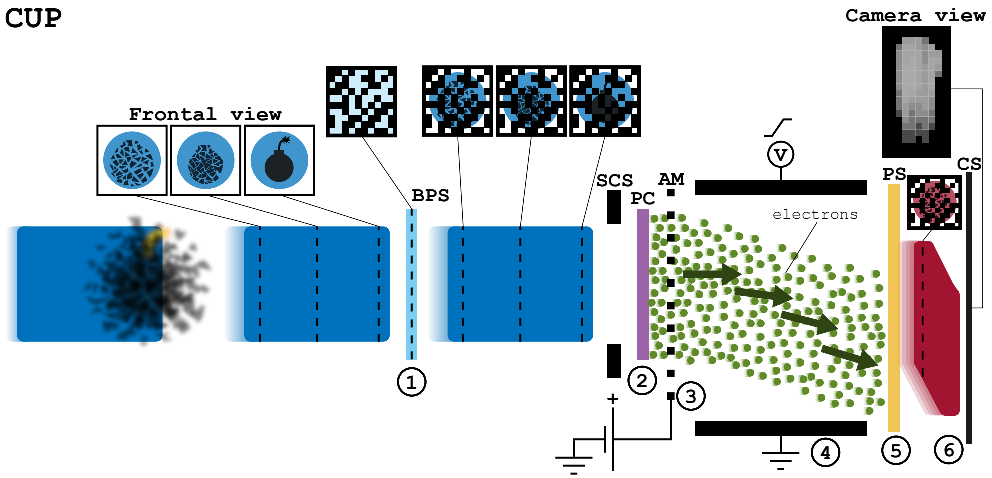


Figure 2.10: Operational principle of CUP. (1) High-intensity light from a dynamic scene is encoded with a random binary pattern. (2) The encoded light enters the streak camera through a fully open slit and is converted into electrons by the photocathode (PC). (3) The electrons are horizontally accelerated by a static voltage applied across an accelerating mesh (AM). (4) The electrons are vertically sheared by a rapidly varying voltage, introducing a temporal dispersion. (5) The temporally sheared electrons are converted back to light by a phosphor screen (PS). (6) The resulting light signal is recorded by the internal camera sensor (CS). BPS: binary pattern screen; SCS: streak camera slit.

Since its introduction, CUP has been developed in a variety of ways. The first being the addition of a second camera to capture an unsheared view of the dynamic scene before the binary pattern is imprinted [114]. The pixel values recorded by this secondary camera set a hard upper limit for the sum of corresponding pixel values in the reconstructed data cube, thereby reducing the number of possible solutions to the compressive sensing problem and improving reconstruction accuracy.

Another significant advancement involved imprinting the binary pattern using a digital micromirror device (DMD) [114]. The DMD deflects light either left or right, and by recording both complementary deflections with the streak camera, the system reduces undersampling and achieves more accurate reconstructions.

Further development introduced active illumination of the dynamic scene with a train of temporally stretched and chirped laser pulses [115]. By incorporating a diffraction grating that shears light horizontally according to wavelength before it enters the streak camera, temporal shearing was achieved in both the y- and x-directions, resulting in a higher effective frame rate.

CUP has also been adapted for specialized applications. For example, it has been combined with dark-field microscopy to enhance sensitivity to light deflected by transparent objects [116], and integrated with a Mach-Zehnder interferometer to achieve true phase sensitivity [117].

In the first CUP demonstration, a frame rate of 50 Gfps ($\Delta T = 20$ ps) was reported, with a temporal resolution measured at 74 ps. The spatial resolution was approximately 3 mm (0.3 lp/mm), with slightly reduced performance in the spatiotemporal shearing direction (the *y*-direction), and a field of view of $50 \times 50\,\mathrm{mm}^2$. However, the standout feature was the sequence depth, which reached an impressive 350 frames.

Four years later, advancements in CUP achieved a frame rate of 10 Tfps ($\Delta T = 100$ fs) with a temporal resolution of 580 fs, maintaining the same sequence depth of 350 frames. This significant increase in frame rate was attributed to the use of a faster streak camera. Further progress was made with the introduction of active illumination, leading to a frame rate of 70 Tfps ($\Delta T = 14$ fs) over 980 frames and a temporal resolution of 240 fs in 2020 [115]. In 2019, a frame rate of 219 Tfps ($\Delta T = 4.6$ fs) was demonstrated, achieving a temporal resolution of 108 fs over 230 frames [118].

Spatial resolution for these higher frame rates was not explicitly reported, but for the 70 Tfps case, it appears to have been slightly over 100 µm, similar to the best explicitly reported value [119]. This value was measured as 109 µm in the *y*-direction and 99 µm in the *x*-direction, based on the apparent broadening of a known-sized object. The observed anisotropy in spatial resolution was attributed to the shearing effect in the *y*-direction.

CUP has been applied to investigate a variety of phenomena, including the refraction, reflection, and propagation of short laser pulses in scattering media in 2D [90, 120, 87] and 3D [85]; plasma dynamics [85]; the propagation of photonic Mach cones [121]; the laser-induced Kerr effect in crystals [116, 115, 118]; the propagation of action potentials in axons [117]; laser induced shock wave propagation in water [116]; and the propagation of laser-induced filaments in glass [118].

One notable limitation of CUP is the lack of a simple connection between the optical system and the resulting images. For most other ultrafast imaging techniques, each light pulse, imprinted with image information, directly corresponds to a single image. Even in cases where a continuous light pulse carries image information, it is usually intuitive to understand how different segments of the pulse are separated and processed to produce an image sequence. For CUP, however, this direct relationship is obscured by the non-intuitive inner workings of compressive sensing (CS) algorithms. While this complexity can lead to skepticism about the integrity of the technique, it should only highlight the importance of calibrating and verifying novel imaging methods against known reference data.

The most significant technical challenge for CUP likely lies in its reliance on a streak camera, rather than the CS reconstruction itself. Within the streak camera, photons are converted to electrons, which inherently differ in their initial velocities. This results in non-uniform shearing for electrons corresponding to the same instant of the data cube, ultimately degrading spatial resolution. Additionally, electron-electron repulsion causes smearing in both time and space, with this issue intensifying as signal levels increase due to higher electron densities. This introduces an undesired trade-off between signal strength and image quality. As summarized by Kim et al., streak cameras *"trade spatial and temporal resolutions as well as dynamic range for higher frame rates"* [116]. While it has been speculated that spatial resolutions as fine as 1 µm could be achieved with the use of more magnifying optics [115], this has yet to be demonstrated. Yet another challenge with using a streak camera is the temporal jitter inherent to all electronic devices, which can make it hard to consistently capture ultrafast transient events without relying on relatively long total exposure times. This limitation is largely mitigated in all-optical techniques, which avoid the timing fluctuations associated with electronic systems.

Another issue lies in the discrepancy between reported and effective frame rates. Reported frame rates are often significantly higher than those supported by the measured temporal resolutions, effectively reducing the practical frame rate. For instance, in the 70 Tfps and 219 Tfps demonstrations, the effective frame rates were approximately 8 and 12 times lower, respectively, when accounting for oversampling. This mismatch is likely due to limitations in the CS reconstruction process, which enforces one frame to be reconstructed for each detector row. However, even if the resulting data were down-sampled to better match the temporal resolution, the reconstructed data cube would still contain significantly more frames than other ultrafast imaging techniques, and CUP's effective frame rate would remain among the highest in the field.

3.4 Spatial frequency division multiplexing

Spatial Frequency Division Multiplexing (SFDM) ensures that frames are separated in the Fourier domain rather than in real space. In this approach, subsequent frames can fully overlap on the detector so that each sensor pixel collects information from all frames. Despite this complete overlap, individual frames remain distinguishable because each probe pulse is encoded with a unique spatial frequency—much like assigning each radio channel its own carrier frequency. By "locking in" to the correct frequency in Fourier space, the corresponding frame can be reconstructed even though it shares the same physical detector area as all the other frames.

Our technique, Frequency Recognition Algorithm for Multiple Exposures (FRAME), operates according to this principle, but it will be discussed in greater detail in Chapter 3.

3.4.1 Single-shot Sequential Holographic Imaging.

The first demonstration of ultrafast SFDM was published in 2006 by Wang et al. [100]. This paper may even represent the earliest instance of ultrafast single-shot videography. While Wang et al. referred to their approach as "spatial angular multiplexing," they did not assign it a specific name. The term Single-shot Sequential Holographic Imaging (SSSHI) was introduced in a 2019 paper that utilizes the same fundamental principle, illustrated in Fig. 2.11.:

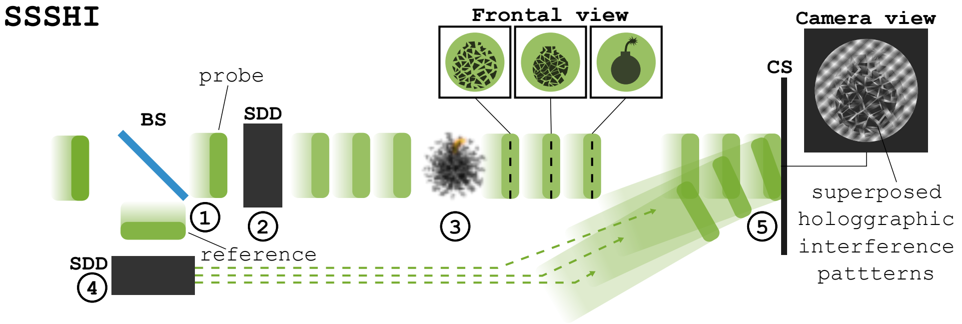


Figure 2.11: Operational principle of SSSHI. (1) An initial short pulse is split into a probe pulse and a reference pulse. (2) The probe pulse is further split and temporally delayed to form a train of probe pulses. (3) The probe pulses interact with the target, encoding information about the dynamic event. (4) The reference pulse is split and delayed to form a train of reference pulses that correspond temporally to the probe pulses. (5) he probe and reference pulses are pairwise overlapped on the detector, forming superposed holographic interference patterns. The orientation and period of these patterns depend on the angle between the corresponding probe pulse and its reference pulse. BS: beam splitter; SDD: split and delay device; CS: camera sensor.

In the 2019 paper a modification to the operational steps of SSSHI was introduced by splitting the initial pulse into sub-pulses early in the process. Each sub-pulse was directed into a dedicated delay arm equipped with a grating that can be rotated around the optical axis. In this setup, the zeroth-order beams generated by the gratings served as the probe pulses, while the +1st-order beams functioned as reference pulses. This approach addressed several challenges. First, it allowed for a common delay stage for each probe-reference pair, greatly simplifying temporal alignment. Second, using gratings instead of beam-splitters to separate the probe and reference pulses enabled perfect spatial overlap between them on the detector. Finally, the rotatable gratings facilitated precise control over the relative angles between the probe and reference pulses, and thus the orientation of their common interference pattern. This was a significant improvement over the 2006 method, which relied on multiple mirrors to control the angles.

The 2006 study demonstrated a frame rate of 3.3 Tfps ($\Delta T = 300$ fs) over three frames, with 50 fs probe pulses and a field of view of approximately 150×310 m^2. The authors claimed a spatial resolution of 10 µm, but this estimation was based solely on the pixel size of the detector. Such a claim does not properly account for optical limitations, as spatial resolution is often determined by optical factors rather than pixel size alone. In SFDM techniques in particular, spatial resolution is further reduced due to the low-pass filtering involved in locking-in to specific frequency components. In contrast, the 2019 study achieved an improved frame rate of 5 Tfps ($\Delta T = 200$ fs) over five frames, using 30 fs probe pulses and a field of view of approximately 165×370 µm^2. From the imaged objects, the spatial resolution in this study can be estimated at approximately 15 µm.

The two studies were similar in their choice of target. The 2006 paper focused on laser-induced ionization of air. While no observable movement was captured across the frames, a clear temporal increase in phase was observed in the ionized region. The 2019 paper investigated laser-induced plasma dynamics, also in air. Here, the plasma signal displayed noticeable movement across the frames, capturing the dynamic evolution of the event.

These techniques share several similarities with FRAME, particularly in their use of spatial frequency encoding to distinguish between frames. However, an important distinction lies in the use of unaffected reference pulses in SSSHI to form the interference patterns. This makes it relatively straightforward to quantify the optical delays experienced by the probe pulses. Nevertheless, this approach introduces certain challenges. The requirement for precise spatial and temporal alignment between probe and reference pulses can complicate the addition of more frames. Furthermore, when using very short probe and reference pulses (e.g. less than 10 fs, corresponding to a spatial extent of less than $3\ \mathrm{m}$), the additional optical path length induced by the target may prevent temporal overlap between the probe and reference pulses, making it impossible to form interference patterns.

Chapter Summary

Even without a precisely defined target, the pursuit of ever-faster videography remains valuable, as pushing the boundaries of temporal and spatial resolution may lead to unexpected, yet groundbreaking, discoveries. However, several well-defined target events exist in ultrafast imaging, particularly in the study of high-intensity light-matter interactions, which require micrometer-scale spatial resolution, femtosecond temporal resolution, and high sensitivity to optical density variations.

In this chapter, I have reviewed a range of ingenious multiplexing methods for ultrafast videography developed by others. In the following chapters, I will present the method I have worked on and contributed to developing.

Chapter 3

Frequency Recognition Algorithm for Multiple Exposures

1 Natural Images

When exploring the world around us through imaging, our primary interest often lies in capturing images of naturally occurring objects—commonly referred to as natural images. In this section some of the key properties of natural images, with a particular focus on their representation in the Fourier domain, will be examined. These properties play a critical role in enabling the optical multiplexing capabilities of Frequency Recognition Algorithm for Multiple Exposure (FRAME), and their exploitation is central to the technique's success.

1.1 A tiny fraction

Natural images represent only a minuscule fraction of all possible images. By natural images, we mean those that could, in theory, depict real-world scenes, particularly scenes involving naturally occurring objects [122]. These are the types of images that most interest physicists, as they reflect the world we aim to investigate. Even images of man-made objects often resemble natural images unless they contain numerous sharp edges or straight lines—features that are relatively rare in the natural world and thus appear ”unnatural.”

The idea that natural images form such a small subset of all possible images might initially seem counterintuitive, given that an infinite number of natural images clearly exist under this definition. However, a thought experiment can help clarify this claim. Imagine digitizing all images into a resolution of 256 by 256 pixels, each with a bit depth of 8, resulting in 256 possible grayscale values per pixel. Now, suppose we start randomly generating such images. What are the odds that one of these randomly created images even remotely resembles a natural scene? Intuitively, we recognize that this probability is vanishingly small. Almost every randomly generated image would appear as meaningless noise. Thus, the proportion of natural images among all possible images must indeed be extraordinarily small [123].

To solidify this understanding, I will soon return to this question and provide an estimation of an upper bound for the proportion of natural images among all possible images.

1.2 A property of natural images

One fundamental property of the natural world—and therefore of images capturing it—is the absence of preferred spatial scales. While we may find objects at the scale of centimeters or meters most relevant to our daily experience, nature itself does not inherently favor any particular size. If we zoom in, shifting perspective from a human scale to that of an ant, previously imperceptible details become prominent, while objects that once seemed large now appear vast. Importantly, the relative proportion of large to small structures remains unchanged across scales.

This scale invariance can be visually observed in Fig. 3.1a, which presents a forest scene at three different magnifications. Details that are indistinguishable at $\times 1$ magnification emerge at $\times 4$ and $\times 16$, yet the overall distribution of large and small features appears consistent across scales. This impression is confirmed when examining the Fourier transforms of these images in Fig. 3.1b. In the frequency domain, high frequencies—corresponding to fine details and sharp edges—are located near the periphery, while low frequencies—representing broad, slowly varying features—are concentrated at the center. Despite differences in zoom level, the Fourier amplitude distributions remain strikingly similar: highest at low frequencies and gradually decreasing towards higher frequencies. This self-similarity is a hallmark of natural images and reflects their scale-invariant frequency distribution [124].

Mathematically, this invariance implies that the average amplitude $\bar{A}$ of Fourier components at a given spatial frequency f follows a power-law relationship:

$$\bar{A} = kf^{\beta}, \tag{3.1}$$

where k and β are real constants. To illustrate this, Fig. 3.1c plots $\bar{A}$ as a function of f for five different zoom levels of the forest scene. Both axes are logarithmic so that the power-law relation in Eq. 3.1 appears as a straight line. The slopes of the fitted lines vary only slightly, from -0.94 to -0.98, closely matching the commonly cited $\beta = -1$ [125, 124, 126]. The data is obtained by grouping the Fourier components into frequency bins, calculating the

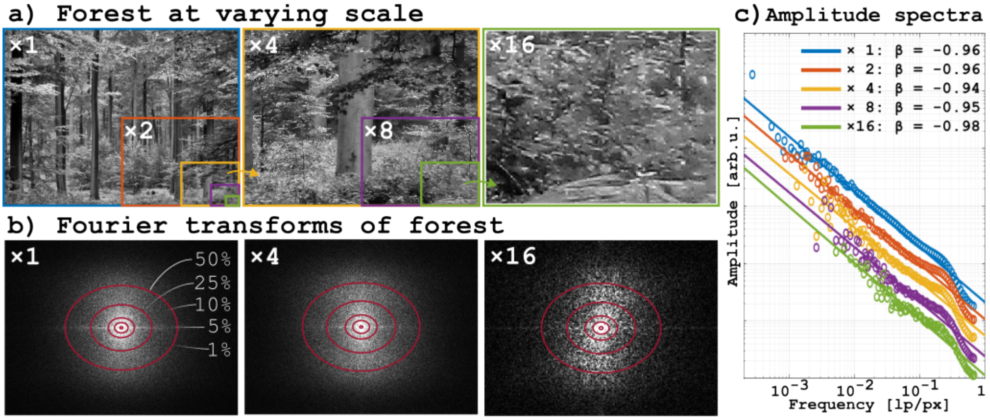


Figure 3.1: Scale invariance of natural images. (a) A forest scene observed at different magnifications, demonstrating the presence of features across a wide range of scales. Credit: Donar Reiskoffer. Licensed under CC BY 3.0. (b) The corresponding Fourier transforms of the images above. The red ellipses indicate regions containing 50%, 25%, 10%, 5%, and 1% of the total spectral power, respectively, illustrating the spatial frequency distribution. (c) The amplitude spectra of the images in (a), plotted as functions of spatial frequency. The similarity across scales highlights the scale-invariant statistical structure of natural images.

mean amplitude and frequency within each bin, and finally fitting a straight line to the logarithms of these averaged quantities.

Empirical studies have confirmed this trend across diverse natural scenes. In a 1992 study, Tolhurst et al. analyzed 135 natural images and found an average exponent of $\beta = -1.2$, indicating that, on average, the Fourier amplitudes decay even more steeply than expected [123]. The study also reported variation across images, with β values ranging from -0.6 to -1.6, highlighting the statistical nature of this scale invariance. [1]

A key observation is that the Fourier transform of natural images is not evenly distributed but is instead heavily concentrated at lower frequencies. This characteristic is fundamental to the compressibility of natural images, as it implies that high-frequency components contribute less to the overall image structure. Consequently, many of these components can be discarded with minimal perceptual degradation [125].

This concept is illustrated in Fig. 3.2, where a beautiful mountain scene is shown after different degrees of low-pass filtering, retaining the most central (lowest frequency) 100%, 10%, 1%, and 0.1% of the Fourier components. As fewer frequency components remain, the image quality visibly degrades, yet the scene remains recognizable down to 1%. Notably, the minimal degradation between the 100% and 10% cases may seem too good to be true—and, in a sense, it is. The reason for this limited perceptual difference is that in the 100% case, the resolution of the printed image is not limited by the underlying image content

---

[1] When repeating the experiment with 500 natural images stolen from Wikipedia I obtained β-values in the exact same range, but with a mean of -1.1.

but rather by the printing process itself. As a result, many of the frequency components discarded when reducing to 10% had no impact on the printed image quality.

This scenario, in which factors other than the imposed bandwidth restriction (low-pass filtering in this case) determine the final signal quality, is ideal for multiplexing. Since multiplexing inherently restricts the available bandwidth for each signal, it is most effective if this restriction does not degrade the final signal quality. Examples of such cases include displaying the image on a low-resolution screen, where high-frequency details would not be resolved regardless, or when the discarded frequency components do not contain meaningful information.

Furthermore, the non-uniform spectral distribution of natural images suggests that the sparsely populated regions of their Fourier transform could, at least in principle, be repurposed to store additional information—such as image data from an entirely different scene. This concept is central to FRAME and will be explored in an upcoming section.

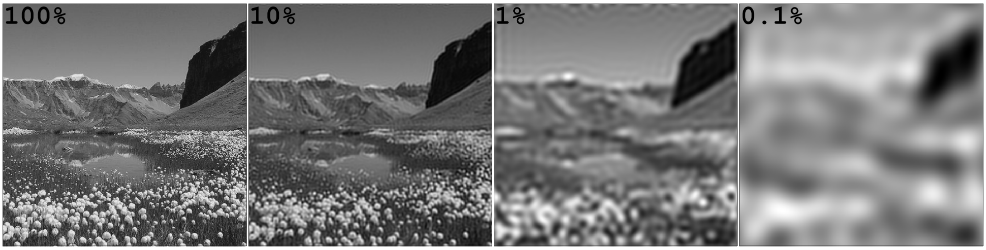


Figure 3.2: The effect of low-pass filtering. The proportion of remaining frequency components is indicated for each image. Credit: Matthias Zepper. Licensed under CC BY 3.0.

1.3 Estimating the proportion of natural images

As previously discussed, natural images constitute only a tiny fraction of all possible images. But just how tiny? To address this question, we leverage the finding by Tolhurst et al. that natural images generally exhibit a power spectrum slope $\beta < -0.6$ [123]. To estimate the proportion of images that meet this criterion, 10^7 random $N \times N$ grayscale images (256 gray levels) were generated for each integer $N \in [4, 20]$ and their corresponding amplitude spectra were calculated. Fig. 3.3a contains examples of an actual natural image, a random image that fulfills the slightly more generous criterion $\beta < -0.5$ and a random image that does not, along with their Fourier transforms. It is important to note that while this criterion does not guarantee that an image could have originated from a real-world scene, it at least identifies candidates that share a key statistical property with natural images.

As expected, the proportion P of images that meet $\beta < -0.5$ and hence are classified as natural decreases rapidly with increasing N. In Fig. 3.3b the fraction starts just below 0.4

for 4 × 4 images and drops to approximately one in a hundred for 10 × 10 images and one in a million for 19 × 19. The data is well described by the exponential fit:

$$P = e^{-0.023N^2 - 0.35N + 0.88}.$$

Extrapolating this fit to $N = 256$ yields a proportion of approximately one in 10^{693}—a number so big it makes astronomical numbers seem insignificantly small. [2] Moreover, since the inclusion criterion for a natural image was rather generous, the true proportion is likely even smaller.

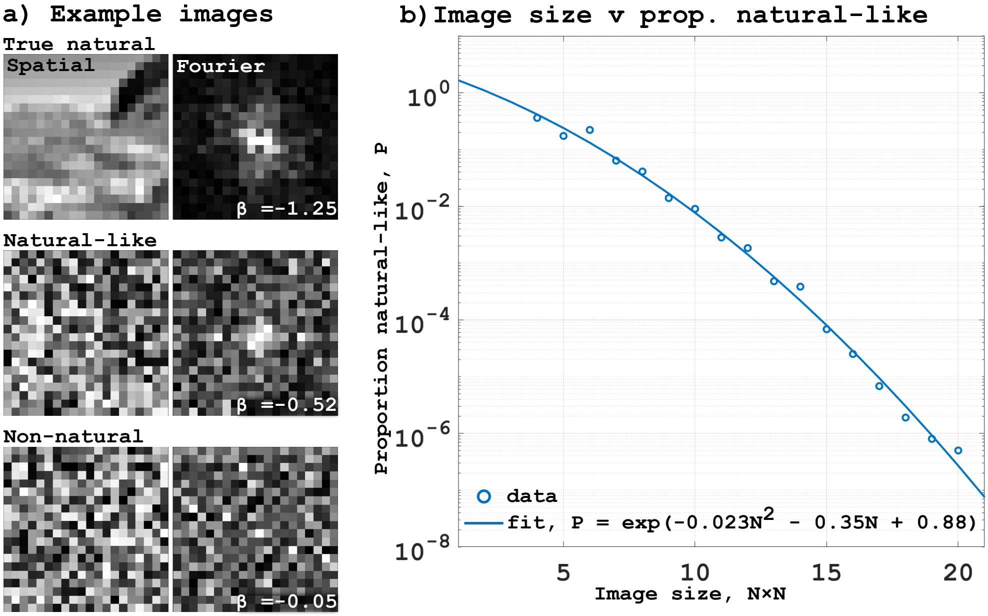


Figure 3.3: Proportion of random images with an amplitude spectrum similar to that of a natural image, as a function of image size. (a) Example 20 × 20 images, including a downsampled version of the mountain scene from Fig. 3.2, a random image with a sufficiently low β to be considered natural-like, and a non-natural image. (b) Proportion of random images fulfilling the criterion $\beta < -0.5$, and thus classified as natural-like, plotted as a function of image size.

2 Modulated illumination

To optically redistribute image information into the underutilized regions of the Fourier transform, modulated illumination is employed. This section explores the principles behind this ap-

---

[2] To put this into perspective, even if all 3.3×10^{80} particles in the observable universe were each generating one random 256 × 256 image every Planck time (10^{-43} s) for the entire age of the universe (4.3×10^{17} s), they would have collectively produced only 1.4×10^{141} images. The probability of generating a single natural image in this scenario would still be less than one in 10^{551}.

proach and examines its effects when used to illuminate a sample.

2.1 The effect of modulated illumination

Having identified the peripheral regions of a natural image's Fourier transform as sparsely populated, these regions present an opportunity for encoding additional image information. The key question then becomes: how can we transfer image information into these underutilized regions? While it is trivial to synthetically extract the central Fourier components of one image and paste them into the periphery of another, such an operation is not optically feasible. Instead, we require a method that is not only theoretically sound but also experimentally realizable. The solution lies in *modulated illumination*, which refers to illumination with a sinusoidally modulated intensity profile, as depicted in Fig. 3.4a. This figure will be used throughout this section to explain the effect modulated illumination has on the perception of a target.

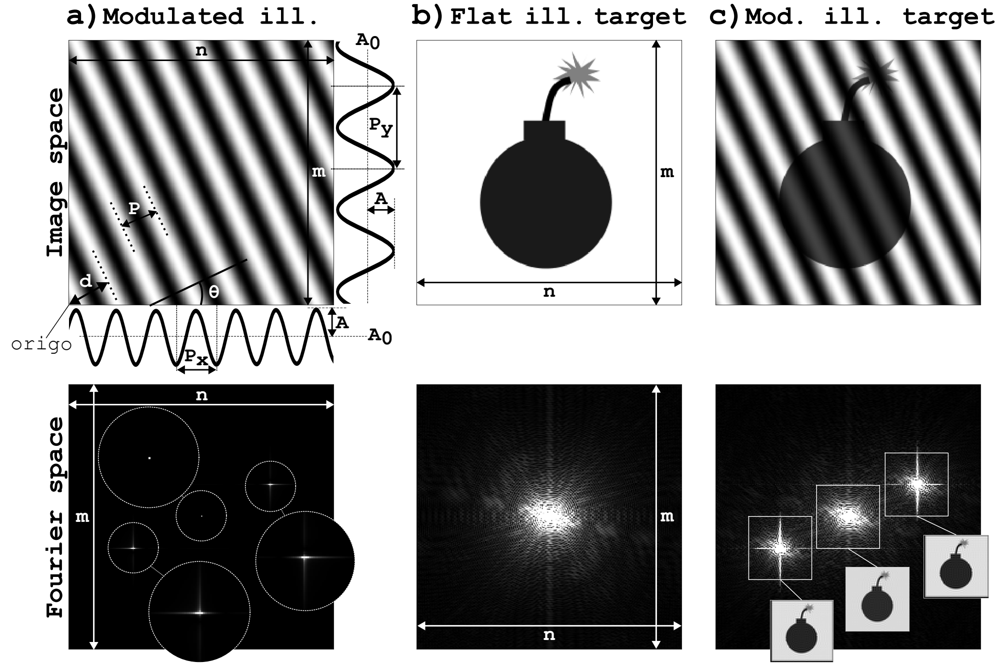


Figure 3.4: Effect of modulated illumination in image and Fourier space. (a) A synthetic image of modulated illumination with indicated parameters, used throughout this section. Three distinct frequency components are visible in its Fourier transform. (b) A synthetic image of a target illuminated from behind with flat (uniform) illumination. Its Fourier transform shows high-amplitude frequency components symmetrically concentrated in the center. (c) The same target as in (b), but now illuminated with the modulated illumination from (a). In the Fourier domain, copies of the original image information are shifted to locations corresponding to the modulation frequency. The low-resolution versions of the target image were reconstructed by applying an inverse Fourier transform to the indicated sections. Note: The period of the modulated illumination is increased tenfold in image space for visibility.

From the detector's perspective, modulated light results in an image similar to the $m \times n$

image in Fig. 3.4a, which is characterized by an amplitude A, an angle θ relative to the x-axis, a period P (or equivalently, spatial frequency $f = 1/P$), and a phase ϕ. The phase is not explicitly stated in the figure, but it can be found by measuring the shortest distance d from the origin to any (preferably the closest) crest and setting

$$\phi = 2\pi \frac{d}{P}.$$

The relations between P and its projections along the coordinate axes are given by $P_\mathrm{x} = P/\cos\theta$ and $P_\mathrm{y} = P/\sin\theta$, respectively. The frequency in vector form is hence

$$\vec{f} = (f_\mathrm{x}, f_\mathrm{y}) = \left(\frac{1}{P_\mathrm{x}}, \frac{1}{P_\mathrm{y}}\right) = \left(\frac{\cos\theta}{P}, \frac{\sin\theta}{P}\right).$$

These values are useful when synthetically constructing 2D modulations to emulate spatially modulated illumination.

Underneath the spatial detector image is a visualization of its Fourier transform, showing the absolute value (or amplitude) of all its frequency components. In this representation, three distinct structures appear: a single central component, located at pixel coordinates $(n/2, m/2)$, proportional to the image's average pixel value, and two cross-like structures at coordinates

$$(u_\pm, v_\pm) = (\frac{n}{2} \pm \frac{n\cos\theta}{P}, \frac{m}{2} \pm \frac{m\sin\theta}{P}) = (\frac{n}{2} \pm nf_\mathrm{x}, \frac{m}{2} \pm mf_\mathrm{y}), \tag{3.2}$$

whose central components are proportional to the amplitude of the sinusoidal modulation. Due to discrete sampling effects from the $m \times n$ pixel detector, the sinusoidal illumination is not perfectly smooth, leading to cross-shaped structures in the Fourier domain rather than single-frequency components.

In Fig. 3.4b, we see a target [3] as imaged by the detector when illuminated from behind with a uniform (flat) intensity profile. The observed image intensity is hence proportional to the transparency of the target scene. Its Fourier transform shows a large symmetric central cluster, containing its constituent frequency components with hightest amplitude.

The image in Fig. 3.4c illustrates the target in Fig. 3.4b when illuminated by the modulated illumination in Fig. 3.4a. In the spatial domain, the effect may not appear particularly striking, but in the Fourier domain, a remarkable transformation occurs: the frequency components of the target image (previously confined to the central region) are systematically shifted to new locations, corresponding to the frequencies of the modulated illumination.

To verify that the displaced components indeed contain the target's image information, one can extract a section (marked by boxes in the figure) of the Fourier transform centered at any

---

[3] A bomb, quite frankly.

of these locations and compute its inverse Fourier transform. The result is a low-resolution version of the target image, as shown in the figure. The smaller the extracted section, the more strongly low-pass filtered the resulting image becomes, leading to reduced spatial resolution. The relationship between filter size and image resolution has previously been illustrated in Fig. 3.2. From the Fourier transform in Fig. 3.4c, it is clear that the section cannot be made arbitrarily large, as increasing its size too much will cause information from neighboring clusters to leak into the reconstructed image. Hence, a certain degree of low-pass filtering is unavoidable when multiplexing images in this manner.

Overall, this demonstrates that we have successfully transferred image information to sparsely populated regions of the Fourier domain using a method that closely mimics what can be reliably implemented in the laboratory.

2.2 Mathematical description of modulated illumination

In the following section, the effect of applying a sinusoidally varying—i.e., modulated—illumination to a target is described mathematically. Although the content may be technical, intuition and illustrative examples will still play a central role.

2.2.1 Sums of sinusoidal terms

In mathematical terms, the duplication and transfer of image information in the Fourier domain can be understood by considering both the modulated illumination and the target images, $I_{\text{mod}}(x, y)$ and $I_{\text{target}}(x, y)$, respectively, as composed of sums of sinusoidal terms terms with periods $P \in [\infty, 2 \text{ px}]$, i.e., frequencies $f \in [0, \frac{1}{2} \text{ lp/px}]$. [4] The phase and amplitude of each of these *spatial terms* corresponds to a pair of *frequency components* in the Fourier domain, as will be demonstrated below.

We begin by expressing the image of the modulated illumination, assuming that the modulation varies from amplitude $A_0 - A$ to $A_0 + A$, with frequency $\vec{f}$ and phase ϕ, as:

$$\begin{aligned} I_{\text{mod}}(\vec{r}) &= A_0 + A\cos(2\pi \cdot \vec{f} \cdot \vec{r} + \phi) \\ &= A_0 + \frac{A}{2} e^{i(2\pi \cdot \vec{f} \cdot \vec{r} + \phi)} + \frac{A}{2} e^{-i(2\pi \cdot \vec{f} \cdot \vec{r} + \phi)} \end{aligned} \tag{3.3}$$

---

[4] The units of the inversely related quantities P and f may appear inconsistent at first glance. A period P corresponds to exactly one line pair (lp), meaning that the unit px/lp would be technically correct. However, we prefer to omit the denominator, as it is implicitly understood. For frequency f, omitting the numerator would feel slightly awkward, so we include "lp" for clarity. That said, using px^{-1} as the unit for frequency would also be valid.

where $\vec{r}$ is the spatial coordinates (x, y) in vector form. This expansion results in three terms, corresponding to the three distinct complex components ν_0, ν_+, and ν_- in the Fourier transform of the modulated illumination (see Fig. 3.4a). At this stage, making this connection without a rigorous proof may seem like a leap of faith. However, it will soon be strengthened, and more formal treatment of this relationship can also be found in, for example, Chapter 4 of [127]. The first term represents the static DC component, while the following two terms correspond to the symmetric off-center components at coordinates $(u_\pm, v_\pm)$, (see Eq. 3.2), with frequencies $\pm\vec{f}$. The complex DC component is

$$\nu_0 = mnA_0,$$

while the off-center components are

$$\nu_\pm = \frac{mnA}{2} e^{\pm i\phi}. \tag{3.4}$$

The absolute value and phase of the frequency components in Eq. 3.4 are hence directly related to the amplitude and phase of the spatial sinusoidal term in Eq. 3.3. This link between a sinusoidal pattern and its Fourier representation is summarized in Table 3.1.

Table 3.1: Fourier components of a sinusoidal $m \times n$ image with spatial frequency $f = 1/P$ at angle θ, amplitude A_{mod}, phase ϕ, and average value A_0. The table shows the corresponding discrete frequency components and their locations in the Fourier transform.

Spatial frequency (f_x, f_y)	Frequency Component	Location in Fourier transform (u, v)
$(0, 0)$	$\nu_0 = mnA_0$	$\left(\frac{n}{2}, \frac{m}{2}\right)$
$\left(\frac{\cos\theta}{P}, \frac{\sin\theta}{P}\right)$	$\nu_+ = \frac{mnA}{2} e^{i\phi}$	$\left(\frac{n}{2} + \frac{n\cos\theta}{P}, \frac{m}{2} + \frac{m\sin\theta}{P}\right)$
$\left(-\frac{\cos\theta}{P}, -\frac{\sin\theta}{P}\right)$	$\nu_- = \frac{mnA}{2} e^{-i\phi}$	$\left(\frac{n}{2} - \frac{n\cos\theta}{P}, \frac{m}{2} - \frac{m\sin\theta}{P}\right)$

Now consider the target $I_{\text{target}}(\vec{r})$ in Fig. 3.4b. Its Fourier transform is much more densely populated, but actually also symmetric, consisting of pairs of frequency components with equal amplitudes and opposite phases. If the proposed connection between an image and its Fourier transform holds, then it should be possible to reconstruct the image by summing sinusoidal terms whose frequencies, amplitudes, and phases match those in its Fourier transform.

In Fig. 3.5, this approach is applied to the Fourier transform of the target in Fig. 3.4b. The reconstruction begins by sequentially adding sinusoidal terms, starting with those with the highest amplitude. As more terms are included, the reconstructed image increasingly resembles the original target. With 5 terms the bomb resembles a blurred circle, with 100 terms the fuse is clearly visible, and with 10,000 terms (out of a maximum of 32,768), the image becomes indistinguishable from the original.

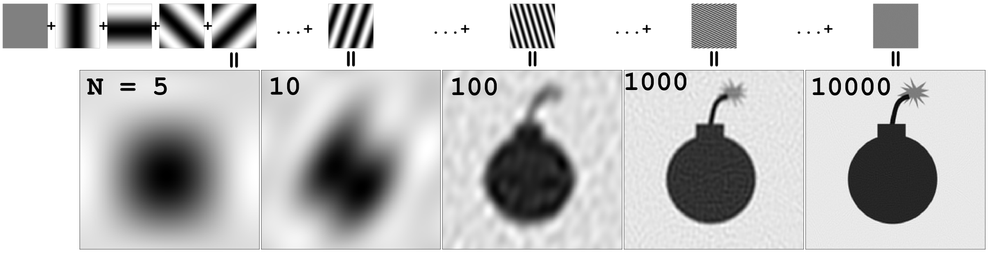


Figure 3.5: Reconstruction of an image by summing sinusoidally varying terms. The amplitude, phase, and frequency of each term are determined from the Fourier transform of the original image. The number of terms used, N, is indicated for each image. The individual terms are displayed above the larger reconstructed images, which are the cumulative sums of these terms.

This result confirms that the target image—or more generally, any digital $m \times n$ image—can indeed be expressed as a finite sum of real-valued sinusoidal terms:

$$I_{\text{target}}(\vec{r}) = \sum_{\text{i}=0}^{N} A_{\text{i}} \cos(2\pi \cdot \vec{f}_{\text{i}} \cdot \vec{r} + \phi_{\text{i}}),$$

where $N \approx \frac{mn}{2}$,[5] $\vec{f}_0 = 0$ and A_0 is the mean pixel value of the image.

2.2.2 Interaction between target and modulated illumination

If we now assume that the target is illuminated from behind by a spatially modulated light pulse (as described by Eq. 3.3), the detected pixel intensity $I_{\text{det}}(\vec{r})$ at each location $\vec{r} = (x, y)$ will be proportional to the product of the uniformly illuminated target's intensity $I_{\text{target}}(\vec{r})$ and the unperturbed modulated illumination $I_{\text{mod}}(\vec{r})$. Assuming that the uniform illumination and the modulated illumination have the same peak intensity, so that the modulation function effectively acts as an envelope to the target image, we have:

$$\begin{aligned}
I_{\text{det}}(\vec{r}) &= I_{\text{mod}}(\vec{r}) \cdot I_{\text{target}}(\vec{r}) \\
&= \Big(A_0 + A\cos(2\pi \cdot \vec{f} \cdot \vec{r} + \phi)\Big) \sum_{\text{i}=0}^{N} A_{\text{i}} \cos(2\pi \cdot \vec{f}_{\text{i}} \cdot \vec{r} + \phi_{\text{i}}) \\
&= A_0 \sum_{\text{i}=0}^{N} A_{\text{i}} \cos(2\pi \cdot \vec{f}_{\text{i}} \cdot \vec{r} + \phi_{\text{i}}) + A \sum_{\text{i}=0}^{N} A_{\text{i}} \cos(2\pi \cdot \vec{f}_{\text{i}} \cdot \vec{r} + \phi_{\text{i}}) \cdot \cos(2\pi \cdot \vec{f} \cdot \vec{r} + \phi),
\end{aligned} \tag{3.5}$$

---

[5] The exact value depends on whether m and/or n is even or odd, but is approximately half of the total number of pixels due to the symmetry of the Fourier transform.

where the first term is an unmodulated version of the target image, and the second term is the *key term*, as it redistributes the target's frequency components by shifting them in the Fourier domain.

To avoid making things overly complicated, we isolate the contribution from the i^{th} term:

$$\begin{aligned} I_{\text{shift,i}}(\vec{r}) &= AA_{\text{i}}\cos(2\pi\cdot\vec{f}_{\text{i}}\cdot\vec{r}+\phi_{\text{i}})\cdot\cos(2\pi\cdot\vec{f}\cdot\vec{r}+\phi) \\ &= \frac{AA_{\text{i}}}{2}\left(\cos\left(2\pi\cdot(\vec{f}_{\text{i}}+\vec{f})\cdot\vec{r}+(\phi_{\text{i}}+\phi)\right)+\cos\left(2\pi\cdot(\vec{f}_{\text{i}}-\vec{f})\cdot\vec{r}+(\phi_{\text{i}}-\phi)\right)\right). \end{aligned} \tag{3.6}$$

What Eq. 3.6 shows is that by illuminating the target with a spatially modulated light pulse (with frequency f, amplitude A and phase ϕ), each of the targets spatial terms (with frequency $\vec{f}_{\text{i}}$, amplitude A_{i} and phase ϕ_{i}) results in two additional terms with new frequencies, phases and amplitudes

$$\begin{aligned} \vec{f}_{\text{i}}^{\mp} &= \vec{f}_{\text{i}} \mp \vec{f}, \\ \phi_{\text{i}}^{\mp} &= \phi_{\text{i}} \mp \phi, \end{aligned}$$

and

$$A_{\text{i}}^{\mp} = \frac{AA_{\text{i}}}{2},$$

respectively. Since each spatial term is represented by two frequency components in the Fourier domain (according to Eq. 3.4) this produces four distinct frequency components in the Fourier domain, which are shifted versions of the two original components:

$$\nu_{\pm}^{\mp} = \frac{mnAA_{\text{i}}}{4}e^{\pm i(\phi_{\text{i}}\mp\phi)},$$

According to Eq. 3.2 these components are located at:

$$(u_{\pm}^{\mp}, v_{\pm}^{\mp}) = \left(\frac{n}{2} \pm n(f_{\text{i},x} \mp f_x), \frac{m}{2} \pm m(f_{\text{i},y} \mp f_y)\right). \tag{3.7}$$

Each frequency component of the uniformly illuminated target image is hence *shifted* by $\pm(nf_x, mf_y)$ in the Fourier domain. Since every frequency component is shifted equally, this process effectively redistributes the entire image information into previously unused frequency regions, making it possible to multiplex multiple images.

3 Implementing FRAME

This section details the practical implementation of FRAME, highlighting the critical steps involved in both the optical setup and the post-processing reconstruction process. The technique is presented as it was when I began my doctoral studies in early 2019.

3.1 The optical setup

Multiplexing images using modulated illumination is the core principle of FRAME, first introduced in 2017 by Ehn et al. [128]. Like single-shot sequential holographic imaging (SSSHI) (discussed in Section 3.4 of Chapter 2), FRAME is a spatial frequency division multiplexing (SFDM) technique, meaning that image information is encoded using distinct spatial frequencies. In FRAME, these spatial frequencies are imprinted as intensity modulations onto the illumination pulse by imaging a grating, whereas in SSSHI, they arise due to interference between probe and reference pulses on the detector. A key distinction is that FRAME does not require interference and can therefore be implemented using either coherent or incoherent light. The operational principle of FRAME can be summarized in five steps, illustrated in Fig. 3.6.

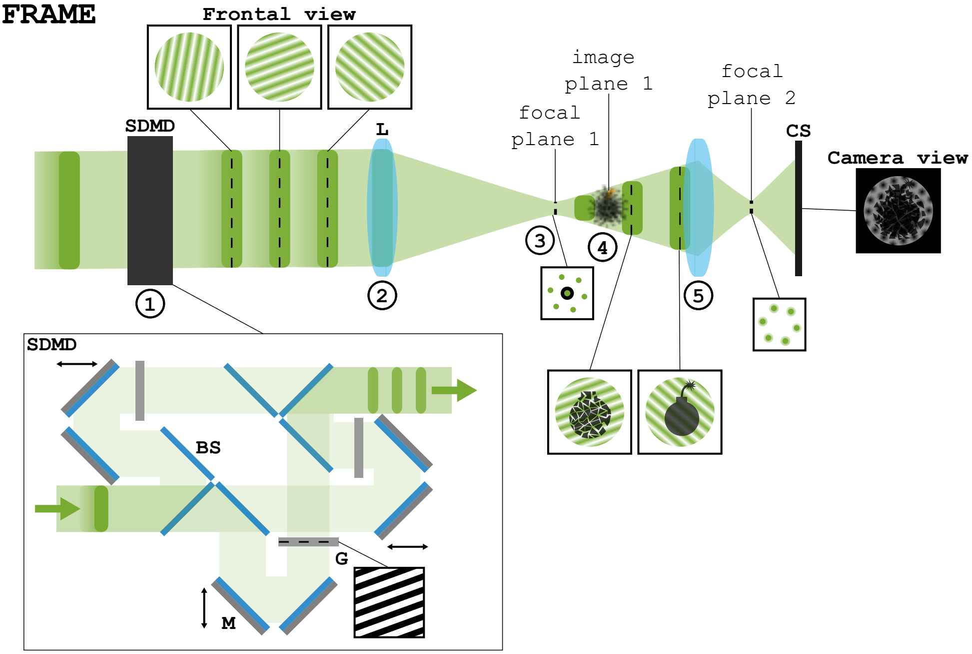


Figure 3.6: Operational principle of FRAME. (1) An initial light pulse is split into multiple sub-pulses, which are individually delayed and pass through gratings with distinct angles and/or frequencies. (2) The resulting train of sub-pulses is directed toward the target, ensuring that the gratings are imaged onto the target plane. (3) If needed, the probe pulses are spatially filtered in the focal plane of the first lens, removing undesired frequency components (e.g. the DC component) to improve the quality of the modulations. (4) The probe pulses interact with the target, encoding the dynamic scene. (5) The probe pulses are directed to the camera sensor, ensuring that both the target and the gratings are imaged onto the sensor. SDMD: split, delay, and modulate device; BS: beam splitter; M: mirror; G: grating; L: lens; CS: camera sensor.

Step 1 is to produce a train of uniquely modulated probe pulses. In principle, these pulses can originate from different sources and be combined using beam splitters to form a pulse train. However, to achieve the highest frame rates, the pulses must originate from the same initial pulse to ensure optical synchronization and avoid temporal jitter associated with electronic timing. Since the primary focus is on ultrafast applications, this description is restricted to laser light from a single common source. In Fig. 3.6, the Split, Delay, and Modulate Device (SDMD) executes this step by splitting the initial pulse with a set of beam splitters, delaying the resulting sub-pulses using orthogonal mirrors on a translation stage to create the desired interframe time ΔT, passing the sub-pulses through gratings with unique orientations to imprint the desired spatial modulation, and eventually recombining the sub-pulses with a mirrored set of beam splitters. It is crucial that all gratings are equidistant from the final recombination beam splitter to ensure that they are all imaged onto the target plane. Additionally, it is beneficial if all sub-pulses to undergo the same number of reflections and transmissions to maintain uniform refraction effects across pulses. [6]

Step 2 is to image the modulated sub-pulses onto the target, which can be achieved using a single lens. If the illumination is incoherent, it is particularly crucial that the image plane of the gratings perfectly match the target plane, as any misalignment would degrade the modulation contrast. For coherent illumination, however, the requirement is less stringent, since sinusoidal modulation naturally emerges when two coherent beams overlap at an angle. When a coherent beam, such as a laser, passes through a grating, diffracted beams emerge at angles dictated by the wavelength of the light and the grating frequency. Imaging the grating is to overlap these refracted beams perfectly, but even imperfect overlap will result in a modulation pattern. Typically, the highest possible spatial frequency is preferred (as long as it remains resolvable by the imaging system), as it allows the highest spatial resolution in the reconstructed images. Consequently, the magnification or de-magnification of of this imaging step should be adjusted so that the size of the pulses matches the size of the target.

Step 3 is an optional step to filter out undesired spatial frequency components in the focal plane (the optical Fourier domain). If a grating is used to generate the modulation, the unfiltered pulse profiles will not be purely sinusoidal; instead, they will exhibit a high-contrast striped pattern with sharp transitions between maximum and minimum intensity. However, by blocking all but the ± 1 diffraction orders in the focal plane, the resulting intensity profile becomes sinusoidal. This effect arises from the way spatial frequency components combine to form the rectangular intensity profile associated with a grating. If unwanted diffraction orders are not blocked, they will show up in the Fourier transform of the multiplexed detector image and possibly result in cross-talk in the reconstructed (de-multiplexed) images.

Step 4 is the interaction between the probe pulses and the transient target. As each tempo-

---

[6] In the figure, the upper left beam splitter serves solely to equalize the number of transmission events for all pulses.

rally distinct pulse passes through the target, its intensity is modulated based on the target's state at that specific moment. Depending on the dynamics of the target, different regions of each probe pulse will be absorbed, transmitted, or otherwise altered, thereby encoding the evolving scene into the sequence of probe pulses.

Step 5 is to image the target onto the detector. Since the gratings have already been imaged onto the target, both the target and the modulation patterns are subsequently imaged onto the camera sensor. To maximize the sensor's efficiency, the probe pulses should preferably fill the entire sensor area, and ideally the full bit depth should be utilized without overexposing a significant number of pixels.

3.2 Reconstruction

Once all the previous steps have been implemented, the camera sensor records an image similar to the camera view in Fig. 3.6, where the temporally distinct illumination pulses have combined into what appears to be a spatially indistinguishable mess. This original image will be called the detector image. [7] In the case of three multiplexed sub-images $I_{\mathrm{i}}(\vec{r})$ (i = 1, 2, 3), each multiplied with a unique modulation $I_{\mathrm{mod,i}}(\vec{r}) = \cos(2\pi\vec{f_{\mathrm{i}}}\vec{r} + \phi_{\mathrm{i}})$ the detector image is

$$I_{\mathrm{det}}(\vec{r}) = I_0(\vec{r}) + \sum_{\mathrm{i}=1}^{3} I_{\mathrm{i}}(\vec{r}) \cdot \cos(2\pi\vec{f_{\mathrm{i}}}\vec{r} + \phi_{\mathrm{i}}), \tag{3.8}$$

where $I_0(\vec{r})$ contains the unmodulated copies of the multiplexed images.

The crucial final step is the reconstruction process, which aims to de-multiplex the image information carried by each individual light pulse, i.e. to retrieve all $I_{\mathrm{i}}(\vec{r})$. This process is illustrated in Fig. 3.7, using a detector image synthetically generated in MATLAB, closely resembling the camera view in Fig. 3.6, as input.

Since all of the necessary steps are illustrated and described in the caption of Fig. 3.7, only the non-trivial steps are discussed in detail below.

Step 2: Constructing the low-pass filter. A suitable low-pass filter (LPF) is constructed based on the Fourier transform of the detector image. Typically, this filter is circularly symmetric, but an elliptical or custom-shaped filter can also be used to match the layout of image information clusters in Fourier space. To avoid excessive cross-talk between images, [8] the filter must not be too large. A reasonable starting point is to set the radius of a circular filter to slightly less than half the distance between the two closest clusters (including the

---

[7] In cases where I have synthetically produced corresponding images—without any involvement of a detector—I will still refer to the resulting image as the detector image.

[8] It is probably not possible to eliminate exactly all cross-talk.

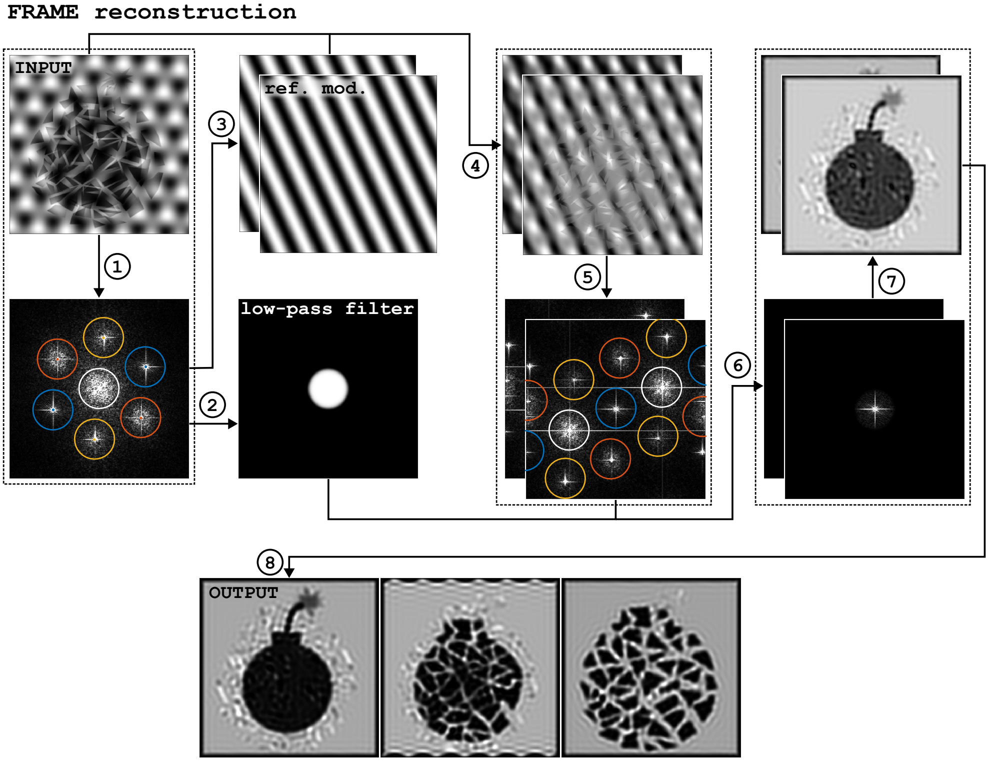


Figure 3.7: Post-processing reconstruction of FRAME in eight steps. (1) The detector (input) image, containing spatial frequency division multiplexed data, is Fourier transformed. (2) A suitable low-pass filter is designed based on the spacing between the image information clusters. (3) Two reference modulations with a phase difference of $\pi/2$ are generated to match the coordinates of the component clusters. (4) Each reference modulation is pixel-wise multiplied with the input image. (5) The resulting images are Fourier transformed, revealing how the previous step shifted the image information clusters, centering the cluster corresponding to the reference modulations. (6) The shifted Fourier transforms are pixel-wise multiplied with the low-pass filter, removing all but the centered image information. (7) The low-pass filtered Fourier transforms are inversely Fourier transformed. (8) The final output image is reconstructed by pixel-wise squaring, summing, and taking the square root of the resulting images. This process is repeated for all multiplexed images. Note: The periods of the modulations are increased tenfold in image space for visibility.

central cluster). Clear signs of cross-talk can be observed in Fig. 3.7, particularly in the still-unexploded bomb image, which is surrounded by noise in the shape of the larger exploded bombs from the adjacent images. This phenomenon is a clear example of *cross-talk*, in which image information from one frame contaminates or overlaps with the information in other frames.

Step 3: Constructing reference modulations. Two reference modulation images $R_{\mathrm{i},1}(\vec{r})$ and $R_{\mathrm{i},2}(\vec{r})$ are generated with the same spatial frequency as the modulation frequency of the probe pulse corresponding to the image being reconstructed. Since the encoding spatial frequency term's phase is unknown, two orthogonal reference modulations (i.e., with a phase difference of $\pi/2$) are required to recover the image. The spatial frequency $\vec{f}_\mathrm{i} = (f_{\mathrm{i},x}, f_{\mathrm{i},y})$ of the reference modulations is determined from the coordinates $(u_\mathrm{i}, v_\mathrm{i})$ of the central component in one of the two corresponding image information cluster as:

$$\vec{f_\mathrm{i}} = \left(\frac{u_\mathrm{i}}{n} - \frac{1}{2}, \frac{v_\mathrm{i}}{m} - \frac{1}{2}\right),$$

according to Eq. 3.2. The reference modulations have an average value of 0 and amplitude of 1, yielding:

$$R_{\mathrm{i},1}(\vec{r}) = \cos(2\pi\vec{f_\mathrm{i}}\vec{r}), \tag{3.9}$$

$$R_{\mathrm{i},2}(\vec{r}) = \cos(2\pi\vec{f_\mathrm{i}}\vec{r} + \pi/2). \tag{3.10}$$

Step 4: Frequency shifting by multiplication. Pixel-wise multiplication of the input image with each reference modulation shifts the image information clusters in the Fourier domain, centering the cluster corresponding to the reference modulations. This effect is analogous to how the original image information was shifted using modulated illumination. A major difference is that the reference modulations have no DC component, meaning that *all* image information is shifted by $\pm\vec{f}$, leaving none at its original location. Multiplying the detector image in Eq. 3.8 with the reference images in Equations 3.9 and 3.10 gives:

$$\begin{aligned} I_\mathrm{det}(\vec{r}) \cdot R_{\mathrm{i},1}(\vec{r}) &= \left(I_0(\vec{r}) + \sum_{\mathrm{j}=1}^{3} I_\mathrm{j}(\vec{r}) \cdot \cos(2\pi\vec{f_\mathrm{j}}\vec{r} + \phi_\mathrm{j}) \right) \cdot \cos(2\pi\vec{f_\mathrm{i}}\vec{r}) \\ &= I_0(\vec{r}) \cos(2\pi\vec{f_\mathrm{i}}\vec{r}) \\ &+ \sum_{\mathrm{j}=1}^{3} \frac{I_\mathrm{j}(\vec{r})}{2} \Big(\cos(2\pi(\vec{f_\mathrm{j}} + \vec{f_\mathrm{i}})\vec{r} + \phi_\mathrm{j}) + \cos(2\pi(\vec{f_\mathrm{j}} - \vec{f_\mathrm{i}})\vec{r} + \phi_\mathrm{j}) \Big). \end{aligned}$$

According to Eq. 3.7, this corresponds to the observed shifts in the Fourier domain.

Step 6: Applying the low-pass filter. The previously constructed low-pass filter (LPF) is applied to the Fourier transform of the products of the detector image and the reference modulation images. This removes all components except those closest to the center (corresponding to $\vec{f_\mathrm{i}} = \vec{f_1}$), yielding:

$$\mathrm{LPF}(I_\mathrm{det}(\vec{r}) \cdot R_{\mathrm{i},1}(\vec{r})) = \frac{\tilde{I}_\mathrm{i}(\vec{r})}{2} \cos(2\pi(\vec{f_\mathrm{i}} - \vec{f_\mathrm{i}})\vec{r} + \phi_\mathrm{i}) = \frac{\tilde{I}_\mathrm{i}(\vec{r})}{2} \cos(\phi_\mathrm{i}), \tag{3.11}$$

$$\mathrm{LPF}(I_\mathrm{det}(\vec{r}) \cdot R_{\mathrm{i},2}(\vec{r})) = \frac{\tilde{I}_\mathrm{i}(\vec{r})}{2} \cos(2\pi(\vec{f_\mathrm{i}} - \vec{f_\mathrm{i}})\vec{r} + \phi_\mathrm{i} - \frac{\pi}{2}) = -\frac{\tilde{I}_\mathrm{i}(\vec{r})}{2} \sin(\phi_\mathrm{i}), \tag{3.12}$$

where $\tilde{I}_\mathrm{i}(\vec{r})$ is a low-pass filtered approximation of the sought multiplexed image $I_\mathrm{i}(\vec{r})$. The effect of the filter is to eliminate all spatial frequencies that are not part of the desired image while preserving the relevant information.

Step 8: Final image reconstruction Since ϕ_i in Eq. 3.12 may vary over $\vec{r}$, the phase components $\cos(\phi_\mathrm{i})$ and $\sin(\phi_\mathrm{i})$ are not necessarily constants that can be removed through normalization. This is why *two* reference modulations are needed for reconstruction. Since $\cos^2(\phi_\mathrm{i}) + \sin^2(\phi_\mathrm{i}) = 1$, the absolute value of the reconstructed image is given by:

$$|I_{\mathrm{rec,i}}(\vec{r})| = |\tilde{I}_\mathrm{i}(\vec{r})| = 2\sqrt{(\mathrm{LPF}(I_\mathrm{det}(\vec{r}) \cdot R_{\mathrm{i},1}(\vec{r})))^2 + (\mathrm{LPF}(I_\mathrm{det}(\vec{r}) \cdot R_{\mathrm{i},2}(\vec{r})))^2}. \qquad (3.13)$$

3.3 The early versions

When I started my work in 2019 FRAME was a very young technique with few publications. In this section I will briefly review the results of these early versions of FRAME.

The initial FRAME studies demonstrated that the technique can be implemented in configurations beyond the one outlined in Fig. 3.6. In [128], the illumination did not originate from behind the target but rather from the side, using modulated laser sheets. This enabled simultaneous probing of a flame at different distances from the camera, allowing for the reconstruction of a 3D snapshot of the flame. In this setup, FRAME recorded not the transmission of light but fluorescence emission from the target, induced by the laser illumination. A similar 3D reconstruction of a gliding arc plasma was also performed by Bao et al. in 2021 [129]. In that study, FRAME's ability to selectively detect modulated light was essential, as it enabled the extraction of the modulated probe pulse-induced fluorescence signal, despite the presence of intense unmodulated plasma emission. Although the paper was published after the start of my doctoral studies, the experimental work predates my own. This study is notable in that the primary objective was not to advance the technique itself, but to use FRAME as a tool to investigate a physically relevant phenomenon—demonstrating the method's practical applicability beyond proof-of-concept experiments.

In 2017, Ehn et al. showcased FRAME in an ultrafast configuration, achieving a frame rate of 5 Tfps ($\Delta T = 200$ fs) over four frames while shadowgraphically imaging the laser-induced Kerr effect in a birefringent crystal [130]. The measured spatial resolution in this setup was approximately 67 µm (15 lp/mm) over a 7×7 mm^2 field of view. The sequence displays how the birefringent effect propagates at the speed of light in barely detectable steps. At the time it was widely regarded as the fastest video sequence ever recorded.

Additionally, FRAME had been demonstrated for multispectral imaging in two separate studies, further highlighting the versatility of the technique. In the study by Dorozynska et al., the simultaneous response from four different fluorophores was probed using three distinct excitation wavelengths [131]. The differing spectral responses of the fluorophores allowed them to be distinguished and identified in a single-shot acquisition. In the work by Li et al., a turbulent flame was illuminated with two different wavelengths, and the resulting laser-induced emissions from combustion species were used to localize OH and CH_2O (formaldehyde) within the flame [132]. This was also achieved in a single-shot, demonstrating the capability of FRAME to perform chemically specific imaging of transient phenomena.

Chapter Summary

Natural images exhibit characteristic spectral properties, with their Fourier transforms heavily concentrated at lower frequencies, making them well-suited for compression and signal encoding. By illuminating a target with modulated light, image information can be redistributed into previously underutilized regions of the Fourier domain. FRAME exploits both the spectral properties of natural images and the effects of modulated illumination to achieve optical multiplexing, enabling the capture of multiple images in a single exposure. This raises several intriguing questions: How many images can be multiplexed using FRAME before the image quality is completely degraded? Could FRAME be combined with microscopy to achieve higher spatial resolution? And could it be integrated with a more sensitive imaging modality than shadowgraphy to enhance its applicability?

Chapter 4

My work

1 Investigating the sequence depth

To fully capture dynamic processes throughout their duration, a large sequence depth is often required. In this section, we investigate whether long sequences are possible and what the effects of increasing the number of multiplexed frames in FRAME are.

1.1 Background

When people are introduced to FRAME, a common question typically arises: "How many frames is it possible to multiplex?" While this question may not have a definitive answer, it remains highly relevant. All else being equal, a video sequence with more frames holds more information about the studied process and is hence superior to a shorter sequence. It would therefore be discouraging if FRAME were fundamentally limited to only a few frames—for instance, four frames, which was the highest demonstrated sequence depth when I began my PhD studies. Thus, the primary goal of my first assigned project was to investigate whether it is possible to drastically increase the sequence depth, enabling video sequences with tens, hundreds, or potentially even a thousand multiplexed frames.

To synthetically demonstrate that FRAME can indeed multiplex a high number of frames is not difficult. This had been done before my work started, and Fig. 4.1 presents my own attempt. After recording a slow-motion video of falling metal flanges, individual frames were extracted, and each thereafter superposed with a synthetically generated unique modulation pattern. The frequencies of the modulation patterns were selected by iteratively determining a (near-)optimal arrangement of 100 frequency components in the Fourier domain (see the bottom image in Fig. 4.1a). Once the component coordinates were set, the corresponding modulation patterns were generated based on the relation between frequency component location and spatial modulation in Table 3.1.

The modulated frames were then pixel-wise summed to synthetically form the detector image in Fig. 4.1a. Looking at the Fourier transform of this image reveals a highly structured placement of frequency components, maximizing their separation while avoiding the strong central DC component. The 100 multiplexed frames were reconstructed using standard FRAME methodology, as described in Section 3.2 of Chapter 3. As demonstrated by the reconstructed images in Fig. 4.1b, it is evident that FRAME can *theoretically* multiplex this many frames, albeit with a certain degree of compromised image quality. Notably, the detector image size is only 1080×1080 pixels. A larger image would allocate more frequency components per multiplexed frame, generally improving image quality if the number of frames is kept constant.

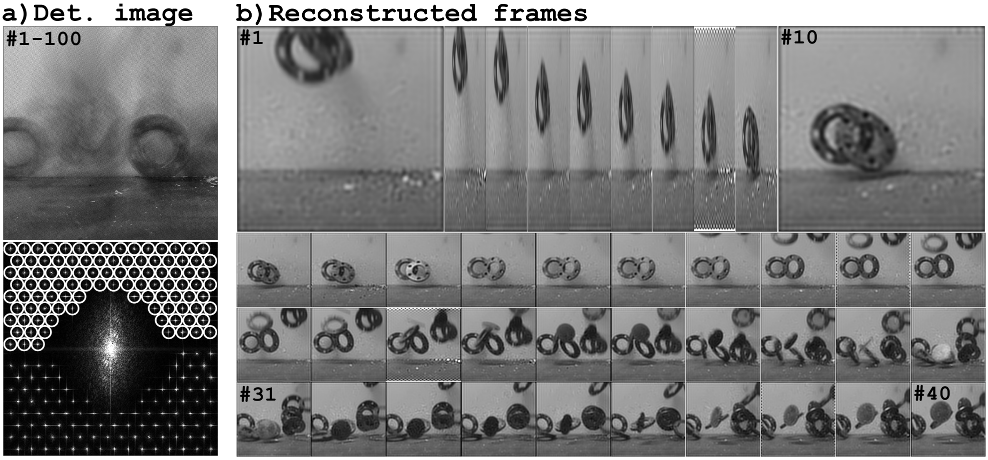


Figure 4.1: Synthetic demonstration of FRAME's multiplexing potential. (a) A synthetically multiplexed image and its Fourier transform. 100 movie frames were superposed with unique modulation patterns and summed to form this image. The white circles in the Fourier transform highlight the upper half of the information clusters used for reconstruction. (b) Reconstructed movie frames from the single detector image in (a). Frames 1 to 40 out of a total of 100 are shown.

Even if a synthetic demonstration can provide an indication of what is possible, it is in no way equivalent to an optical demonstration—which is what ultimately matters. In a synthetic scenario, the modulations are perfect, with no amplitude or frequency deviations across the image. This level of perfection cannot be expected in an optical implementation. For instance, even a slight misalignment of the camera will introduce a skewed observation angle, causing the modulation frequency to vary across the image. Other potential optical challenges include variations in illumination intensity and inconsistencies in modulation amplitude across frames. Therefore, an optical demonstration is not just beneficial but absolutely essential to validate the technique's potential for long sequences.

1.2 Optical setup

To perform the optical demonstration, the setup in Fig. 4.2 was designed and built. It differs significantly from previous FRAME setups (such as the one illustrated in Fig. 3.6), as it does not rely on beam splitters to divide an original probe pulse into multiple consecutive pulses. If the number of pulses N is chosen as a power of 2, the conventional approach would require $2(N-1)$ 50:50 beam splitters to generate the desired pulse train. [1] Additionally, another $2N$ mirrors, arranged in orthogonal pairs on translation stages, would be needed to control the time delay of each pulse.

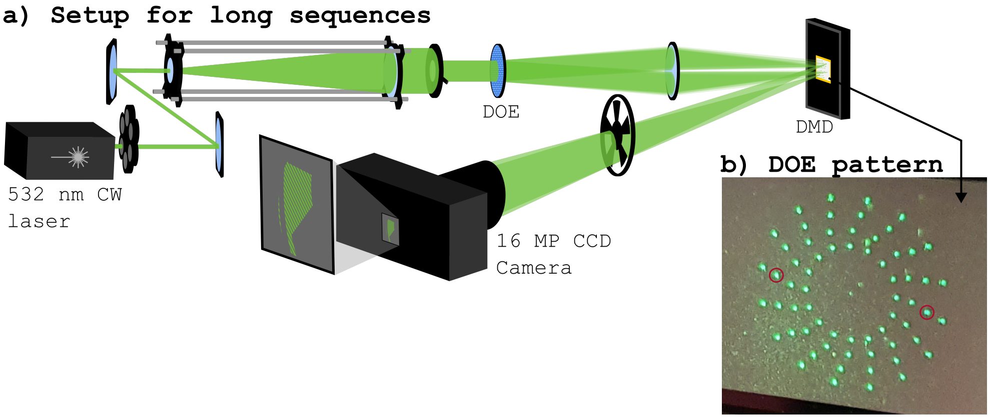


Figure 4.2: Optical setup for long-sequence FRAME. (a) Schematic of the setup used in Paper 1. The key components include the diffractive optical element (DOE), which splits the incoming beam into multiple sub-beams, a lens that images the DOE onto the target, and the digital micromirror device (DMD), which controls the timing of the sub-beams. (b) Photograph of the sub-beams on the DMD, showing the characteristic pattern of the DOE.

In such a setup, each pulse undergoes N beam splitter reflections or transmissions—half of these for splitting and half for recombination. Each recombination results in a 50% intensity loss, leading to a total remaining intensity of $\left(\frac{1}{2}\right)^{N/2}$ relative to the original pulse. For $N = 16$, this type of configuration would require 30 beam splitters, 32 mirrors, and result in an intensity loss of 99.6%. For $N = 128$, the numbers become even more extreme: 254 beam splitters, 256 mirrors, and a remaining intensity proportion of $\sim 5 \cdot 10^{-20}$. While the $N = 16$ case is technically achievable, it would be extremely challenging to align and operate, whereas the $N = 128$ case is entirely unrealistic due to the severe intensity loss. Thus, a different approach is necessary to investigate larger sequence depths.

Instead of using beam splitters, the new setup employs a Diffractive Optical Element (DOE) to split the illumination. A DOE is a thin glass plate with a sub-micrometer-scale surface pattern that causes transmitted coherent light (such as laser light) to interfere constructively only in specific, predefined directions. These diffraction angles can be freely

---

[1] This is likely the wisest choice, if using 50:50 beam splitter.

designed, allowing for the projection of any binary pattern. The pattern used in this setup, shown as a photo in Fig. 4.2b, consists of 64 spots arranged in a spiraling pattern, with four spots per arm across 16 arms. It was designed by my supervisor to mimic a well-structured Fourier domain and the intended use was to only interfere symmetric beam pairs, such as the pair highlighted with red circles in Fig. 4.2b. The Fourier transform of the resulting interference pattern would then match the location of the two interfering beams (except for some scaling factor and an additional DC component). By sequentially imaging all 32 such beam pairs, which are symmetric under 180-degree rotation around the center, their combined interference patterns result in an image whose Fourier transform closely replicates the DOE pattern itself.

The light source in this setup is a continuous wave (CW) laser with a wavelength of $\lambda = 532$ nm. Since the output beam is relatively narrow, it is expanded using a telescope before being cropped by an iris to produce a collimated beam with the desired diameter and a near top-hat intensity profile. This beam is incident on the DOE, which transmits copies of it in 64 unique directions, corresponding to the previously described pattern. As the beam copies propagate away from the DOE, they become increasingly separated. Before this separation becomes too large, they pass through a lens that causes them to start converging.

Near the focal plane of the lens, a Digital Mirror Device (DMD) is positioned. The DMD consists of 1920×1080 micromirrors, each of which can be tilted by 12° either to the left (tilt on) or to the right (tilt off). By carefully controlling the tilt pattern of the DMD, two selected beams at the time can be directed towards the target and eventually to the camera, while the remaining beams are dumped. The target is placed at the plane where the two selected beams perfectly overlap.

If all beams were simultaneously directed towards the target, it would be customary to say that the DOE was imaged onto the target by the lens. However, since only two out of 64 beams are used at a time, it is more appropriate to describe the setup as overlapping beams at the target plane, thereby creating a unique interference pattern. By rapidly shifting the pattern on the DMD, different beams pairs can be selected such that the spatial modulation pattern illuminating the target changes in rapid succession. This leads to the encoding of the target state into temporally distinct modulations. By imaging the target with a camera while it is being exposed to this burst of illumination, a multiplexed detector image is captured, which can then be used to reconstruct the individual frames.

Although the DOE was designed with the intention of only utilizing beams in symmetric pairs, an important realization was made while working with this setup: any two beams can be used, as there is nothing inherently special about a symmetric pair. Any pair will suffice, and the resulting interference pattern will simply be determined by their relative positions, such that the period is dictated by their separation and the orientation by their relative angle.

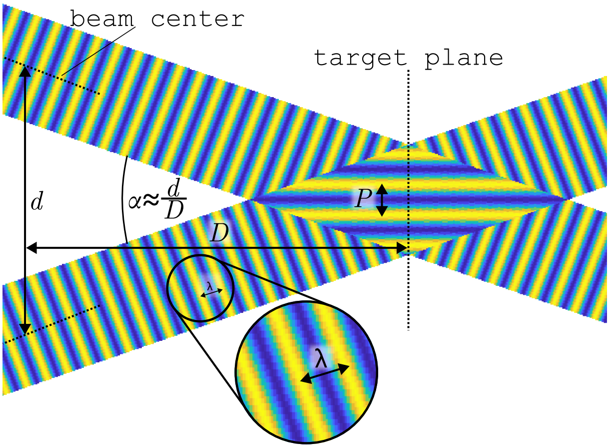


Figure 4.3: An illustration of two overlapping, and thus interfering, beams of coherent light.

To be precise, if the selected beams are separated by a distance $d = \sqrt{d_x^2 + d_y^2}$ and the distance from the DMD to the target, D, is much greater than d, the separation angle between the beams is given by $\alpha \approx \frac{d}{D}$. Using the small-angle approximation $\sin\alpha \approx \alpha$, [2] the period of the interference pattern is:

$$P = \frac{\lambda}{2\sin\alpha} \approx \frac{\lambda}{2\sin(d/D)} \approx \frac{\lambda D}{2d}.$$

This situation is illustrated in Fig. 4.3, where two beams can be seen to overlap and produce an interference pattern at the target plane. [3] The periods P_x and P_y in each direction are found by replacing d with the respective separations d_x and d_y. According to Table 3.1, if an $m \times n$ image of the interference pattern is recorded the corresponding frequency components will be located at

$$(u_\pm, v_\pm) = \left(\frac{n}{2} \pm \frac{n}{D\lambda} d_x, \frac{m}{2} \pm \frac{m}{D\lambda} d_y\right), \tag{4.1}$$

in the Fourier transform of the image, i.e., at positions proportional to the beam separation $\vec{d} = (d_x, d_y)$ at the DMD. Thus, any unique combination of d_x and d_y results in a distinct interference pattern with uniquely located Fourier clusters.

---

[2] ”There is an international law, saying that when you are doing physics for more than a half hour, you must use at least one small angle approximation.” - 3Blue1Brown[133]

[3] The image in question is produced using the following MATLAB code:

```
I = abs((exp(1i*2*pi*(imrotate(ones(100,1)*(1:1000), +alpha/2)/lambda)) + ...
exp(1i*2*pi*(imrotate(ones(100,1)*(1:1000), -alpha/2)/lambda))).^2);
```

The number of possible modulation patterns that can be generated with this setup, given the DOE in question, is therefore limited by the number of beam spot pairs with unique separations. Assuming the DOE pattern is *perfectly* symmetrical, this number is exactly 1024, with the explanation for this provided in Paper 1. This significantly extends the sequence depth beyond the initially intended 32 frames, allowing for the investigation of much longer multiplexed video sequences.

1.2.1 Finding beam positions

To generate an image sequence with $N \leq 1024$, one must simply select N unique beam combinations and sequentially project the corresponding on-off patterns onto the DMD. A prerequisite for this procedure is that the exact locations of the beams on the DMD are known. However, determining these positions is a tedious task, and even minor disturbances[4] can necessitate recalibration. The best approach I found was to use a live script that allowed tilt-on discs to be moved around on the DMD. While observing the DMD through a mobile phone camera, the position and size of the discs were adjusted until all spots were perfectly covered.

1.2.2 Selecting beams

Selecting beam pairs involves finding the N pairs that result in the best distribution of Fourier clusters. Specifically, an optimal distribution minimizes cross-talk between frames by maximizing the spread of Fourier clusters, while also avoiding interference from the strong central DC component. An example of a well-distributed frequency layout is shown in Fig. 4.1a, where most of the available area is utilized while maintaining a safe distance to the DC component.

Finding *any* such distribution is easy, but finding the *optimal* distribution is non-trivial, when the possible Fourier cluster locations are pre-determined and the relative power [5] between individual cluster and the DC component is only known in theory. Assuming that the positions of the beam spots on the DMD have been determined, the best approach I found was as to:

1. Identify the symmetric twin of each beam spot, by finding the spot closest to the symmetric location.
2. Construct all 1024 possible beam pairs and calculate their characterizing separation

---

[4] For example, going home for some sleep.

[5] Power is here used to describe the summed absolute values of a group of frequency components.

distances $\vec{d_i} = (d_{i,x}, d_{i,y})$, which can be thought of as Fourier cluster locations, according to Eq. 4.1.

3. Remove all clusters that are closer to the central DC component than a set threshold.
4. Iteratively identify the smallest distance between any two clusters and out of the two, remove the one closest to any *other* cluster.
5. Either repeat step 4 until exactly N clusters remain or until slightly more than N remains and remove the last ones manually.

With manual removal some obviously ill placed (e.g. directly above or to the side of the DC component) clusters can be avoided.

1.3 Results and further discussion

In this section, a selection of results from Paper 1 is presented. The associated discussions reflect an improved understanding compared to when the study was first conducted.

1.3.1 Dynamic targets

Using the setup and methods described above, two main dynamic targets were imaged: a computer fan spinning at 3000 rpm and a water injection event. For the computer fan, sequence depths of 32, 64, 128, 256, 512, and 1024 frames were acquired. The Fourier transforms of the captured multiplexed images for each case can be found in figure 5 of Paper 1, while examples of the reconstructed frames are shown in figure 4. Typical frames for different sequence depths are shown in Fig. 4.4a.

From these results, it is evident that FRAME can be used with sequence depths in the order of hundreds of frames. However, it is also clear that the image quality of individual frames decreases as the number of frames increases. This trend was subjectively quantified by classifying all extracted frames on a five-degree scale from *good* (5) to *useless* (1). The average ratings for each sequence depth were 4.91, 4.78, 4.08, 3.4, 2.71, and 1.68, respectively.

While quantifying something as subjective as perceived image quality might seem unconventional, it is arguably reasonable. Image quality is a complex attribute that is difficult to reduce to a single numerical value, yet humans have an inherent ability to recognize good and bad image quality at a glance. Therefore, a subjective classification may still serve as a useful metric in evaluating the practical limitations of FRAME at high sequence depths.

The results indicate that the limit for this particular setup is somewhere between 512 and 1024 frames. When the number of frames was doubled from 512 to 1024, the average rating

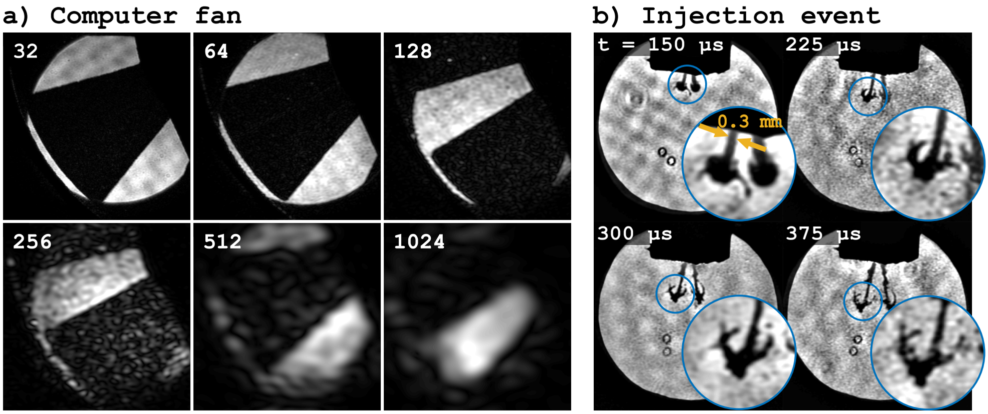


Figure 4.4: Selected results from the investigation of longer FRAME sequences. (a) Typical reconstructed frames for different sequence depths, with sequence depth indicated in the upper left corners. (b) A portion of a 32-frame sequence of a water injection event, acquired using FRAME.

dropped sharply, and the proportion of useless frames—where no meaningful information about the position or shape of the spinning fan blade could be extracted—increased from 0% to 36%.

Another dynamic target, in the form of a water injection event was recorded at 13.3 kfps ($\Delta T = 75$ µs) with a sequence depth of 32 frames. The resulting image series is shown in figure 9 of Paper 1, with a portion also displayed in Fig. 4.4b. Since the refresh rate of the DMD imposed a limit on the achievable frame rate, and the most relevant part of the injection event occurred within the first 2 ms, there was no motivation to use a longer sequence depth. [6] Unlike the spinning fan, this event is stochastic, and the dynamic object occupies only a fraction of the entire image, increasing the risk of overexposure, making it a potentially more challenging target. Although the images show some artifacts, the shadow of the injected fluid can be readily observed throughout all reconstructed frames, and features as small as 0.2 mm (see figure 9 of Paper 1) can be traced over multiple frames, despite the configuration not being microscopic.

1.3.2 Stationary target

To make a truly objective quantification of the relation between sequence depth and image quality, a stationary star (36 spokes) target, was imaged with varying sequence depths. This target, being highly unnatural with razor-sharp edges in every direction, is poorly suited for multiplexing using FRAME. An underlying assumption for FRAME is that the targets are natural, producing images with weak high-frequency components, as discussed extensively in Sections 1.2 and 2.1 of Chapter 3. As seen in the experimental image in Fig. 4.5a, the

---

[6] $2\ \text{ms}/75\ \mu\text{s} \approx 27$

Fourier transform of a star target image exhibits strong features in the shape of straight lines extending to the resolution limit of the optical system. Each of these lines is the result of a perpendicular sharp edge in the detector image. Additionally, since the target is stationary, the lines from all multiplexed sub-images add up in the Fourier domain, making their collective absolute value comparable to that of the multiplexed signals. Consequently, interference from the DC component becomes a major issue in this case. Next to the experimental image are two images of similar star targets that were synthetically constructed and multiplexed ($N = 24$) using the FRAME methodology, while writing this thesis. When the synthetic star target remains stationary, it exhibits the same lines in the Fourier domain. However, when made to rotate, these lines disappear. The drastic effect this has on the reconstructed image quality is evident in Figure 4.5b, where the image reconstructed from the stationary star shows significantly worse quality than its rotating counterpart. This is further highlighted in Figure 4.5c, where the Modulation Transfer Function [7] (MTF) is much more suppressed for the stationary star (black line) compared to the rotating star (white line).

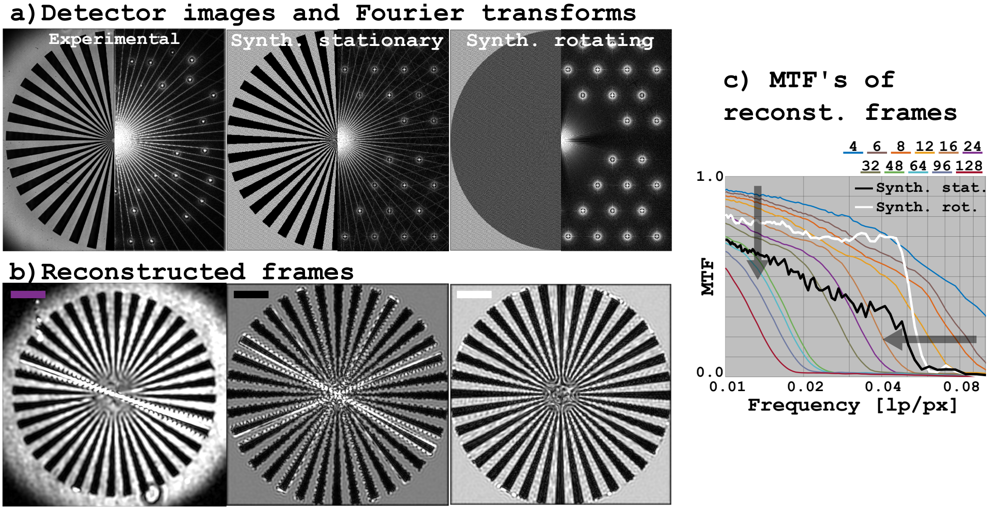


Figure 4.5: Synthetic and experimental star targets. (a) Synthetic and experimental detector images, each containing 24 multiplexed sub-images, along with their Fourier transforms. The synthetic images feature either a rotating or stationary star target, while the experimental image contains a stationary target. (b) Frames reconstructed from the detector images in (a). (c) Modulation transfer functions (MTFs) of the reconstructed frames in (b), as well as MTFs from experimentally obtained frames at varying sequence depths. The horizontal arrow indicates a decrease in cut-off frequency with increasing sequence depth, while the vertical arrow denotes a general reduction in contrast.

With that said, we did not prioritize finding a way to make the physical star target rotate, hence, in Paper I a stationary target was used to quantify the MTF for different sequence depths. Unfortunately, in analysing the effect of sequence depth on the MTF, I focused only on the cut-off frequency (arbitrarily defined as the frequency where the contrast decreases

---

[7] Essentially, the contrast displayed in the star target images as a function of distance to the center, and hence as a function of spatial frequency.

below 0.10). The results showed a monotonically decreasing relation between the cut-off frequency and sequence depth, as illustrated in figure 8a of Paper 1. While this finding is valid, it overlooks an important aspect: not only does the cut-off frequency decrease as the number of frames increases, but the contrast of frequencies *below* the cut-off also diminishes. When re-analyzing some of the reconstructed images and plotting their MTF's in the same graph in Fig. 4.5c this trend becomes very clear. The horizontal arrow indicates a decrease in cut-off frequency, which is directly connected to increased degree of low-pass filtering, while the vertical arrow indicates that the contrast of lower frequencies also decreases as the number of multiplexed frames goes up. What this implies is that image quality deteriorates in at least two distinct ways when using FRAME, rather than just one, as was more or less explicitly assumed in the paper.

1.3.3 Mechanisms for reduced image quality

In an effort to identify the mechanisms contributing to the deterioration of image quality in FRAME, a model was developed during this thesis work, allowing different effects to be introduced to a synthetic star target. Since a dynamic object is typically the intended target for FRAME, the star target was made to rotate while being synthetically multiplexed using the FRAME methodology. The goal of the model is to replicate both the appearance of an individual multiplexed frame and the modulation transfer function (MTF) of that frame.

In the model, the ground truth frames are constructed with a contrast of 0.9 between black and white spokes, along with low levels of Poisson-distributed background noise. These frames are multiplexed using standard (synthetic) FRAME methodology, and the i^{th} reconstructed multiplexed frame serves as a reference. To emulate the observed effects on the multiplexed frame, the corresponding ground truth is subjected to the following three modifications:

1. **Low-pass filtering:** A filter identical to the one used during FRAME reconstruction is applied. The size of the filter is dictated by the distance between Fourier clusters.
2. **Reduction of dynamic range:** The dynamic range is reduced according to the number of multiplexed frames. The reduction factor is determined as the ratio between the maximum values of the ground truth image and the detector image containing all multiplexed frames.
3. **Introduction of artificial cross-talk:** Random noise is introduced in the Fourier domain, with one noise term added for every other Fourier cluster, including the DC component. The phase of the noise is random, while its amplitude follows the power law of natural images in Eq. 3.1, where k and β are determined from the ground truth image. Since the power law is intended for natural images, and the star target

is highly non-natural, k was adjusted by a factor of 1.3 for better resemblance with the multiplexed frame. The amplitude of the DC component is assumed to be $2N$ times greater than the amplitude of a single Fourier cluster.

Each of these modifications reduces the image quality of the star target, as seen both in the resulting images in Fig. 4.6a and their corresponding MTFs in Fig. 4.6b. The effect of low-pass filtering (yellow line) alone primarily results in the suppression of contrast for frequencies beyond the filter's cut-off frequency. While this effect is also seen in the reconstructed image, it does not fully explain the degradation in image quality observed in the multiplexed frame (light blue line).

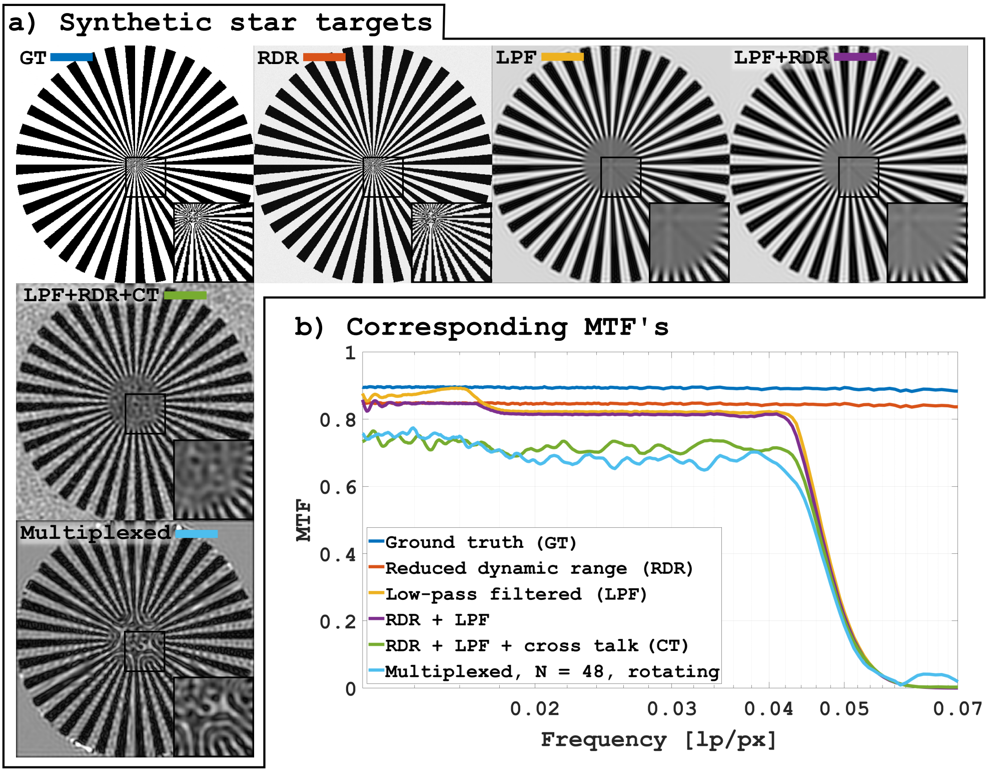


Figure 4.6: Investigation of image quality reduction mechanisms in FRAME. (a) Synthetic star targets subjected to various effects (except for the ground truth) to simulate the degradation observed in images multiplexed and reconstructed using the FRAME methodology. (b) Modulation transfer functions (MTFs) corresponding to the star targets in (a).

Reducing the dynamic range, while keeping the noise constant, to account for less dynamic range being allocated for individual frames when recording multiples frames with FRAME, results in an even reduction of contrast, across all frequencies (red line). The combination of low-pass filtering and reduced dynamic range does not compound linearly; instead, the low-pass filtering smooths out noise introduced by the dynamic range reduction. The result is an MTF that for each frequency is close to the minimum MTF of either modification in isolation (purple line). Further introducing synthetic cross-talk results in the characteristic wobbly appearance seen in the reconstructed multiplexed frame. This effect arises when high-frequency noise, added e.g. through cross-talk, is low-pass filtered, producing larger (than pure noise) random structures. The addition of cross-talk also further suppresses the MTF (green line) at low frequencies, making it more closely resemble that of the multiplexed frame.

By adjusting the number of multiplexed frames, noise levels, and bit depth, the relative appearance of the synthetically multiplexed and reconstructed frame and the frames with synthetically reduced image quality changes. However, the general conclusion remains: low-pass filtering alone does not fully account for the reduction in image quality associated with FRAME and the wobbly appearance of the multiplexed frame is only replicated when cross-talk is simulated by adding noise in the Fourier domain.

2 Shlieren imaging with FRAME

Due to the nature of ultrafast dynamic phenomena, a technique with high sensitivity to refractive index gradients is likely to have an advantage when recording such events. In this section, we investigate whether FRAME can be combined with schlieren imaging to enhance its sensitivity to these effects.

2.1 Background

As detailed in Section 1.2 of Chapter 2, several potential targets for ultrafast FRAME are transparent. To visualize them with good contrast, a technique more sensitive than shadowgraphy would be, if not mandatory, at least highly beneficial. To investigate whether FRAME can be integrated with the more sensitive schlieren techniques, we planned to follow the traditional spatial filtering approach of classical schlieren imaging.

2.1.1 A path towards schlieren imaging

In essence, this approach involves placing a light-blocking object at the focal spot of an imaging system. When a transparent target causes light rays to deviate from their initial paths, they may start hitting (or missing) the block, thereby decreasing (or increasing) the intensity of the corresponding part of the image. In a FRAME setup, the focus of the illuminating light consists of several distinct spots, corresponding to the spatial modulations of the illuminating pulses. Our plan [8] was thus to design a custom spatial filter that precisely matched the focal spot pattern of the FRAME setup and insert it in the focal plane. This would ensure that only light deviated by the target would pass through the spatial filter and contribute to the superposed interference patterns at the camera sensor.

This approach required that the setup produced exactly the intended modulations; otherwise, the spot pattern would not align with the spatial filter. The setup we intended to use splits an initial laser pulse into two and later recombines the sub-pulses at an angle, producing an interference pattern. By utilizing acousto-optical devices (AODs), the recombination angle can be rapidly adjusted with great precision, enabling the generation of a rapid burst of unique modulations.

Testing this setup with a burst of 38 modulations, I found that the resulting Fourier clusters were almost perfectly spaced, with a maximum deviation of only 2% from the intended separation, as illustrated in Fig. 4.7a. This was an encouraging result, suggesting that with only minor manual adjustments, all focal spots could likely be made to match the spatial filter. Consequently, the filter shown in Fig. 4.7b was designed, ordered, and delivered, with experiments scheduled to commence.

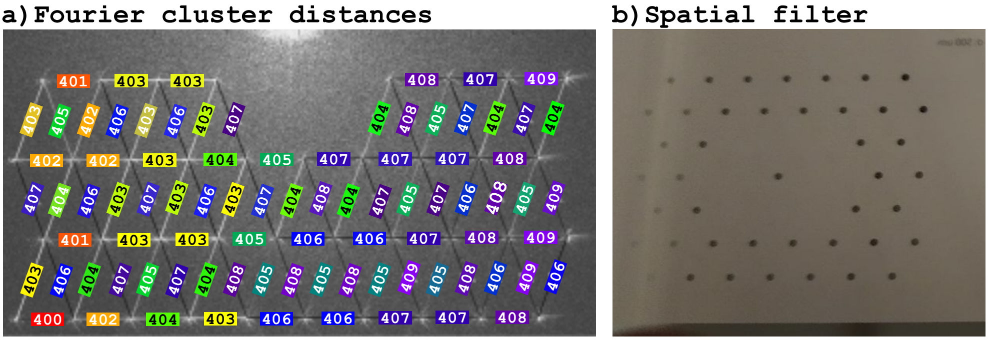


Figure 4.7: Preparations for schlieren experiments with spatial filtering. (a) Measured distances between Fourier clusters, verifying modulation accuracy. (b) The designed spatial filter for schlieren experiments, intended to match the modulation pattern in the focal plane.

At this moment, my supervisor posted the photo of his iPad at a barbecue in Fig. 4.8a, seemingly thrilled about the implications. I must admit that my reaction was somewhat underwhelming, as my understanding of the significance was, at that point, non-existent.

2.1.2 Changing path

It turns out that the wrinkles in the modulation pattern in Fig. 4.8b result from light being deflected on its way from the iPad screen to the camera sensor. What induces this deflection

---

[8] ”Deviced on a sunny afternoon in Elias garden.” - Dr. V. Kornienko.

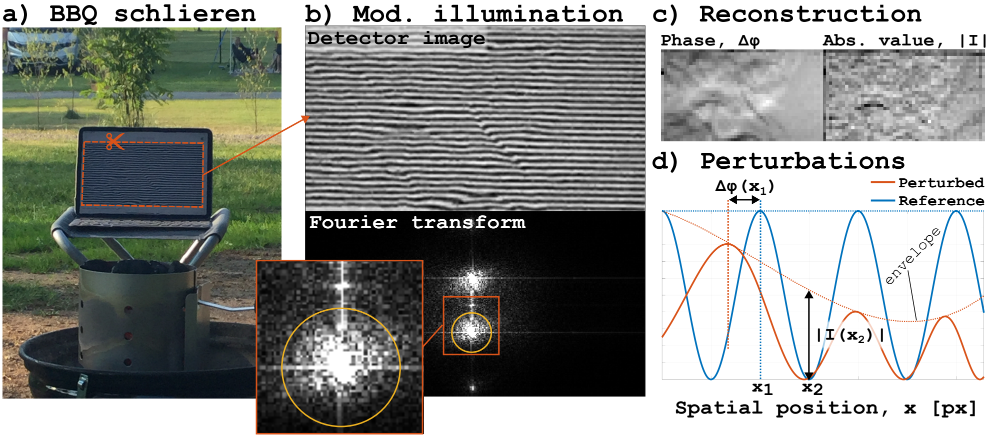


Figure 4.8: Modulated illumination in front of a hot grill, revealing refractive index gradients. (a) An image of an iPad displaying a horizontal striped pattern in front of a hot grill. (b) A zoom-in on the screen, showing apparent wrinkles in the uniform modulation. The Fourier transform reveals that the off-center Fourier clusters associated with the modulation frequency are much more spread out than the DC component. (c) The phase and absolute value of the complex image reconstructed from the part of the Fourier transform within the yellow circle in (b). While the absolute value image is difficult to interpret, the phase image resembles a plume of hot air and is assumed to reveal associated refractive index gradients. (d) A 1D illustration of how perturbations, in the form of a phase change ($\Delta\phi$) and the introduction of an envelope, affect a sinusoidal modulation. These perturbations can be calculated from a perturbed modulation and its unperturbed counterpart. In this spontaneous BBQ imaging scenario, no true reference exists.

is invisible to the naked eye—slight variations in the refractive index of the hot air rising from the grill. If the screen were uniformly lit, (almost) no signal from the hot air would be detected, as it would not induce any change in the perceived image. Similarly, if all refractive index gradients were strictly horizontal, the corresponding light deflections would also be horizontal and therefore remain invisible, merely shifting black areas into other black areas and white areas into other white areas. Conversely, if the screen displayed a static random pattern, the deflection could only be detected by comparing the current image to a reference image taken under undisturbed conditions. [9] This is the principle behind background-oriented schlieren (BOS), where differences in the appearance of an object's background are revealed by comparison to a reference image captured under undisturbed conditions [94]. The only reason we can see the effect in the single image in Fig. 4.8b is that we know how the modulation pattern *should* look. We assume that the lines should be straight, allowing us to clearly perceive the distortions—something that would be impossible with either a uniform or a random background. An early version of BOS actually utilized a structured background like this, calling the technique *synthetic schlieren* [93].

Just realizing the effect that refractive index gradients have on the perception of the modula-

---

[9] Mirages, such as the shimmering air above hot asphalt, are analogous to the rising hot air from the grill. If you try to capture these mirages with a camera, you will see nothing, regardless of how clearly you perceive them with the naked eye. The only reason you can see the mirage is that your brain continuously compares current visual input to previous images, detecting differences over time. With just a single photograph, no difference can be detected, and the mirage remains invisible.

tion patterns we use for multiplexing purposes in FRAME led us to abandon the initial plan of using spatial filters for schlieren imaging. The reasoning was that with such a striking effect, valuable information must surely be encoded in the corresponding Fourier clusters, and our challenge was merely to retrieve and interpret it correctly. That something indeed has been encoded can be seen in the Fourier transform in Fig. 4.8b, where two large Fourier clusters appear at locations corresponding to the modulation frequency.

A particularly interesting observation is that these off-center clusters are significantly larger and contain more power than the DC component. This is not the case when encoding is achieved through modulation of amplitude, as seen in Fig. 4.1a and Fig. 4.5a. In those cases, at least half of the power remains in the DC component, with one-fourth allocated to each of the Fourier clusters, in accordance with Equations 3.5 and 3.6. [10]

The reason for this difference is that the (primary) effect a transparent target has on a modulation

$$I_{\text{mod, i}}(\vec{r}) = A_0 + A\cos\left(2\pi\vec{f}_{\text{i}}\cdot\vec{r} + \phi_{\text{i}}\right)$$

is to introduce a perturbation in the form of a spatially varying phase shift $\Delta\phi(\vec{r})$, such that

$$\phi_{\text{i}} = \phi_{\text{i}}(\vec{r}) = \phi_{i,0} + \Delta\phi(\vec{r}). \tag{4.2}$$

Unlike amplitude perturbations, which act as envelopes (see Eq. 3.5), an added phase shift does not interfere with the spatial DC *term* A_0 and is therefore not encoded in or near the spectral DC *component*. These two types of perturbations are illustrated in Fig. 4.8d.

The absolute value of the i^{th} target image $I_{\text{i}}(\vec{r})$ can be retrieved from a multiplexed detector image $I_{\text{det}}(\vec{r})$ as described in Sec. 3.2. From the expressions in Equations 3.11 and 3.12, it is straightforward to also extract the phase of the i^{th} modulation pattern:

$$\phi_{\text{i}} = \tan^{-1}\left(\frac{\sin\phi_{\text{i}}}{\cos\phi_{\text{i}}}\right) = \tan^{-1}\left(\frac{\text{LPF}(I_{\text{det}}(\vec{r})\cdot R_{\text{i},1}(\vec{r}))}{\text{LPF}(I_{\text{det}}(\vec{r})\cdot R_{\text{i},2}(\vec{r}))}\right), \tag{4.3}$$

where the reference modulations $R_{\text{i},1}(\vec{r})$ and $R_{\text{i},2}(\vec{r})$ are defined in Equations 3.9 and 3.10. By first making a reference measurement where $\phi_{\text{i}} = \phi_{i,0}(\vec{r})$ (which may also vary with $\vec{r}$), the phase perturbation associated with the target, $\Delta\phi_{\text{i}}(\vec{r})$, can be determined using Eq. 4.2. From the detector image in Fig. 4.8b, the complex target image

$$I_{rec}(\vec{r}) = |I_{rec}(\vec{r})|e^{i\Delta\phi(\vec{r})}$$

was thus reconstructed using Equations 3.13 and 4.3, while working on the this thesis. The phase and absolute value of this complex image are shown separately in Fig. 4.8c. While the

---

[10] Only if the amplitude of the modulation is maximized to equal the mean value of the modulation, i.e., if $A_0 = A$, otherwise even more power remains in the DC component.

absolute value mostly resembles noise, the phase is strikingly similar to a typical schlieren image, [11] and clearly reveals a plume of hot air rising from the grill, albeit at low resolution.

2.2 Optical setup and results

The setup used to investigate whether FRAME can be combined with schlieren imaging without the additional step of adding spatial filters is schematically illustrated in Fig. 4.9. It is built around an illumination unit that delivers microsecond modulated incoherent light pulses, which were used to trans-illuminate a Bunsen burner flame. The illumination unit was designed to produce four pulses, but one LED was broken, so only three could be used. [12] Further details about the setup can be found in Paper III.

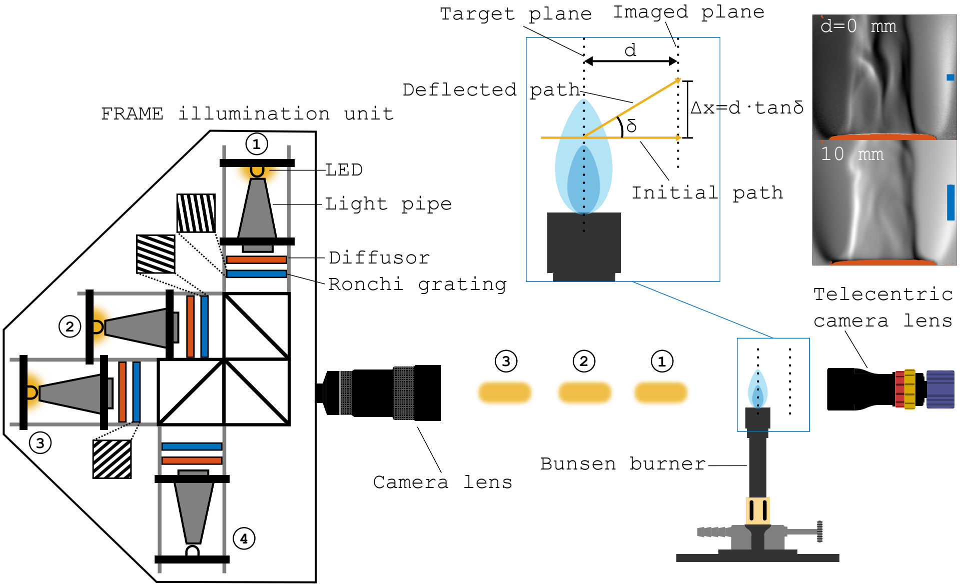


Figure 4.9: Schematic illustration of the setup used to investigate the integration of FRAME with schlieren imaging. The core of the setup is the FRAME illumination unit, which delivers modulated nanosecond pulses of incoherent white light. The relationship between the target-dependent deflection angle δ, the distance d between the target and imaging planes, and the measured displacement Δx is illustrated. Insets show the resulting phase images for $d = 0$ mm and $d = 10$ mm. The dynamic range is significantly greater for $d = 10$ mm, as indicated by the larger blue bar.

All schlieren imaging relies on a target-induced angular deflection $\delta(\vec{r})$ in the paths of the light rays. This deflection results in a spatial displacement

$$\Delta x = d \tan \delta,$$

---

[11] The internet is full of those.

[12] This was during the Great Pandemic, when global supply chains lay in ruin, and the estimated delivery time for a single replacement LED was a full year.

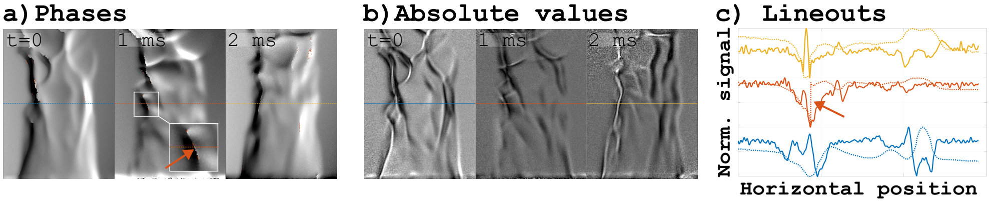


Figure 4.10: Comparison of (a) phase and (b) absolute value images of a turbulent flame. (c) Normalized lineouts from the central row of each image. The lines of the phase images are dashed. The red arrows point at the same discontinuity in (a) and (c).

where d is the distance from the point of deflection to the plane imaged by the camera. When the displacement affects a spatial modulation, it is equally relevant to describe it in terms of an induced phase change

$$\Delta\phi = \frac{2\pi\Delta x}{P},$$

where P is the period of the modulation. The greater the distance d, the more sensitive the system becomes to deflections, as an identical deflection results in a larger measurable displacement (or phase change). However, increasing d also causes the target to become increasingly out of focus, leading to a more blurred image.

This expected trend of increased dynamic range with larger d was confirmed in an experiment where the flame was incrementally translated in 10 mm steps. The resulting image series showed that as d increased, the dynamic range improved, but so did the occurrence of regions where the phase could not be properly reconstructed, due to blurring of the image. Two images from that series are included as insets in Fig. 4.9. The size of the dynamic ranges, which in this context is defined as the difference between the maximum and minimum of the observed phase shifts, are indicated by the length of the blue lines in the images.

A typical multiplexed time resolved sequence from these experiments is displayed in Fig. 4.10, which contains the phase (a) and absolute value (b) images of a turbulent flame, acquired at 1 kfps. The physical interpretation of the phase images is straightforward: a refractive index gradient at the target induces a deflection in the illumination, which leads to a displacement in the recorded image. Due to the modulated nature of the illumination, this displacement appears as a wrinkle in the detector image and can be quantified as the phase change of the modulation.

The physical interpretation of the absolute value images is less clear. One possible explanation is that the refractive index gradients in the flame not only deflect light but also act as a lens. In some areas, light is focused, increasing the modulation amplitude, while in others, light is defocused, decreasing the modulation amplitude. This effect is generally very weak compared to, for example, the light emitted by the flame, but since the modulated illumination allows it to be isolated, it can still be observed using FRAME. My opinion is that the absolute value images can be considered shadowgraphic images with complete background removal.

Regarding the actual appearance of the images, both provide a 3D impression. However, the structures revealed in the absolute value images appear unrealistic for a flame, displaying thin strands trailing the flame's edges. In contrast, the phase images offer a representation that appears more physically plausible. An additional advantage of the phase images is their improved signal-to-noise ratio (SNR). This is evident in the lineouts shown in Fig. 4.10c, which correspond to the central row of each image. The lineouts from the phase images are noticeably smoother, as a consequence of a greater SNR.

2.3 Estimating the sensitivity

To investigate the sensitivity of this schlieren-like technique, I developed a simple model where a small displacement could be introduced to a synthetic modulation pattern. To ensure the model was as realistic as possible, the distribution of pixel values was matched to real data by adjusting the signal and background levels, as well as the overall noise level. The resulting synthetic noisy data closely resembled real data, as demonstrated by the close agreement between the histograms in Fig. 4.11a. In Fig. 4.11b, a central region in the detector images has been shifted by the indicated sub-pixel amount. While no visible change can be observed in the three images with the smallest displacements, the phase images reveal detectable shifts for displacements larger than 0.02 pixels, demonstrating that even with real-world noise levels, the technique is highly sensitive. The signal-to-noise ratio (SNR) of the phase images was calculated based on the size of the detected phase shift and the general noise levels in the phase images. By varying the displacement, the SNR changed accordingly, and Fig. 4.11c shows the median SNR plotted as a function of the modulation period. The results indicate that a modulation period close to 3 pixels yields the highest sensitivity. Coincidentally, a period of 3.4 pixels was used throughout the experiments, aligning well with the observed optimal sensitivity. Further detail can be found in Paper III.

An issue with the sensitivity investigation in Paper III, as described above, is that it was not applied to multiplexed data, which is the standard use case for FRAME. To address this, the model was updated to incorporate multiplexed data, while working on this thesis. Another limitation of the previous model was that the displacement was introduced abruptly, in a highly unnatural manner. To improve realism, the displacement is now applied in the form of a circular super-Gaussian profile.

The example images in Fig. 4.12a show multiplexed ($N = 4$) detector images along with one of the reconstructed phase images. The relationship between the median SNR and the

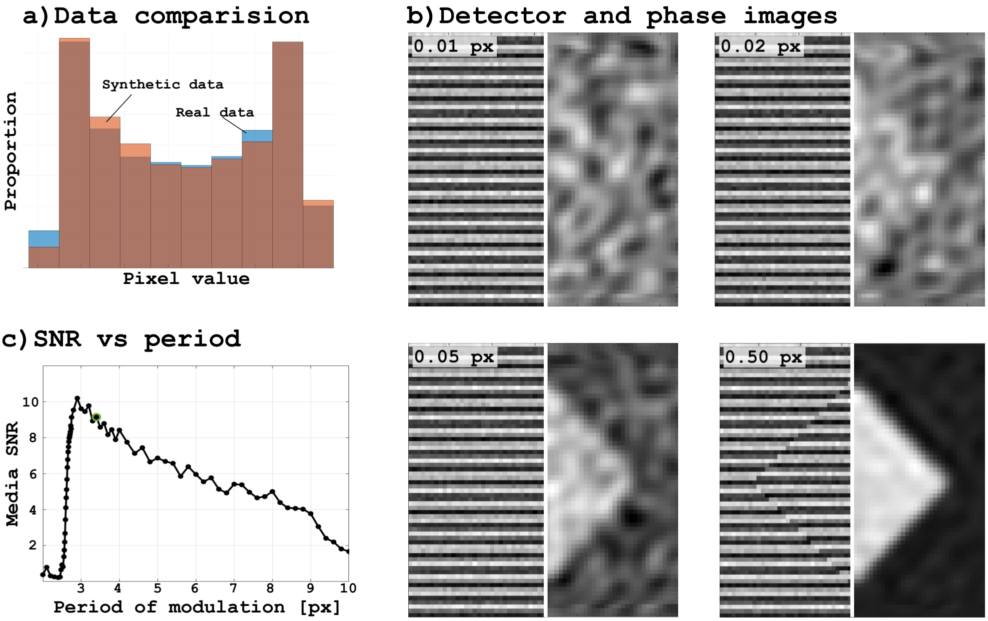


Figure 4.11: Investigation of schlieren sensitivity. (a) A histogram comparison between real data and synthetic data to ensure realistic synthetic noise levels. (b) Detector images (left) with a central section displaced by the indicated number of pixels and the corresponding reconstructed phase images (right). (c) The median SNR in phase images similar to those in (b) as a function of the modulation period. For each period, the SNR was computed over a range of displacements to obtain the median SNR.

modulation period again exhibits a peak near a period of 3 pixels, similar to the previous investigation. Since the updated model differs from the previous one, the SNR values cannot be directly compared.

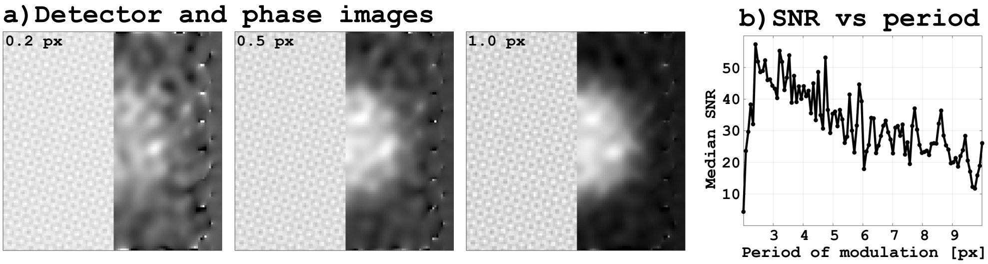


Figure 4.12: Investigation of FRAME multiplexed schlieren sensitivity. (b) Detector images (left) with a central section displaced by the indicated number of pixels and one of the corresponding reconstructed phase images (right). (c) The mean SNR in phase images similar to those in (b) as a function of the modulation period. For each period, the SNR was computed over a range of displacements to obtain the median SNR.

Considering the optimal modulation period for FRAME, it is often assumed that a shorter period is always preferable, as it theoretically allows for higher spatial resolution in the reconstructed images. This aligns with the Nyquist-Shannon sampling theorem, which states that the highest observable frequency is half of the sampling frequency [86]. However, we show here that the period (the inverse of frequency) should not necessarily be minimized, as doing so beyond a certain point decreases the sensitivity to phase changes. This is a somewhat surprising finding that will be relevant for designing future FRAME systems.

3 Finally going ultrafast

The main purpose and motivation of this thesis have been technological development to enable imaging of ultrafast dynamic events with high spatial and temporal resolution. In this section, we examine the experiments where FRAME was applied to ultrafast physics, specifically laser-induced plasma events in air.

Having discovered that FRAME inherently detects phase changes, the natural next step was to apply this capability to an ultrafast transparent target. Hopefully, this marks the beginning of a shift from research primarily focused on the technique itself and its development toward a greater emphasis on the physics that FRAME enables us to study. The goal was (and remains) to move from exploring what FRAME *can do* to focusing on what FRAME allows us to *observe*.

3.1 Designing the setup

When operating at ultrafast timescales, it is not feasible to temporally control the probe pulses using a DMD (as in Paper I) or to use different light sources (as in Paper III). Instead, the probe pulses must originate from the same initial pulse and be controlled purely optically to eliminate any temporal jitter associated with electronic components. [13] This requirement led us to adopt the well-established approach of using beam splitters to divide a single pulse into a pulse train.

In the design process, we aimed to optimize the design by both improving upon previous experiments and preparing for future studies. Thus, when designing the setup, our primary goal was to make the core split, delay, and modulate device (SDMD) compact, easy to set up, and portable so that it could be seamlessly integrated into future experiments. The intent was to build a versatile workhorse that could serve reliably across different experiments and laboratories over an extended period.

Meeting these demands naturally required significant effort in designing the SDMD. Early in the process, we decided that the SDMD should be primarily 3D-printed, with fixed slots for the necessary optical components. This approach would ensure minimal ambiguity

---

[13] Our laser has an incredibly low electronic jitter of $\sim$ 6 ps, but it is still not low enough.

during assembly and reduce the need for extensive alignment, making the setup both robust and user-friendly.

Requiring that the device be 3D-printed meant that its footprint was ultimately limited by the size of the printing bed (330 mm × 220 mm). The small footprint of the device meant that the number of distinct optical paths, and hence the number of probe pulses and ultimately the number of video frames, had to be restricted to four. It would simply not be feasible to double the number of paths within such a small area. [14]

A key component in this type of setup is the translation stage, which allows the optical path lengths to be adjusted so that the timing of the pulses can be controlled. Striving for a 1 fs temporal jitter implies that these stages must have a translational resolution of less than 1 µm. They also needed to be small enough to fit within the limited area. The model that best suited these criteria was the one shown in Fig. 4.13d, the Q-545 from Physik Instrumente. It is driven by a piezo motor and offers a well-rounded 6 nm resolution over a total travel range of 13 mm, within a footprint of 45 mm × 48 mm. Since the light travels back and forth over the translation stage, 13 mm translates to a total temporal adjustability of

$$(13 \times 2)\ \mathrm{mm}/c \approx 9\ \mathrm{ps}$$

per pulse, which is more than sufficient for the intended purposes.

During the procurement of these translation stages, I mistakenly mixed up the maximum loads for vertical and horizontal placement, which led me to wrongly assume that the maximum load was 100 g instead of 500 g. With the smallest commercially available adjustable mirror mounts weighing around 60 g each, and the setup requiring two per translation stage, I decided to design and 3D-print my own. Initially, I attempted static mirror holders, naively believing that the 3D-printing process would be precise enough to eliminate the need for adjustability. That futile dream was soon shattered, and I had to design a holder with adjustable pitch (up-down) and yaw (left-right) angles.

The final mirror holder design is displayed in Fig. 4.13b. It consists of a stationary back part that attaches to a translation stage platform at a fixed angle, and a movable front part that holds the mirror. I opted to use thin strands of rubber, cut from training resistance bands, to hold the two parts together. The rubber bands pull the front part towards the back while still allowing it to rotate. Two standard hex screws protruding through the back adjusts the exact rotational angles. The rubber bands were secured using wedges, as seen in Figures 4.13b and c. The photo shows the complete construction, including mirrors, screws, and broken rubber bands. Since rubber degrades relatively quickly under tension [134, 135], I more than once returned from vacation to find the setup in need of renovation and realignment due to shattered rubber bands.

---

[14] Doubling is typically how the number of pulses is increased when using beam splitters.

Figure 4.13: Renderings and photos of the split, delay and modulate device (SDMD) designed for ultrafast FRAME. (a) 3D renderings of the SDMD. A single laser path enters the unit, where it is split and delayed into four paths, each consisting of two paths separated by a small angle. In the laser path view, each path is shown in a distinct color. For visualization purposes, the separation angle within path pairs has been exaggerated from 1.91° to 6°. (b) Components designed to hold optical elements, including a platform accommodating two mirror holders. Each mirror holder consists of a front and a back part, enabling independent adjustment of the pitch (up-down) and yaw (left-right) angles of the mirror. The wedges are designed to secure rubber bands that hold the parts together. (c) A photo of the delay unit, showing an orthogonal pair of mirrors in their holders mounted on a translation stage. The orange rubber bands have broken due to being under stress for an extended period of time. (d) The translation stage around which the unit was designed.

Aside from the mirror holders, the other optical holders (for DOEs and beam splitters) were static and proved to work well. The finished device consists of a base plate with slots for the four translation stages, as well as for the DOE and beam splitter holders. A platform with slots for the mirror holders is mounted on top of the translation stages. A 3D rendering of the complete device is found in Fig. 4.13a, which also includes a top-down view of the four laser paths. In Fig. 4.14 an schematic illustration show how the SDMD was integrated in the final setup, used throughout Paper v. In this setup the device was used to orthogonally trans-illuminate a laser induced ultrafast event, revealing the associated refractive index variations.

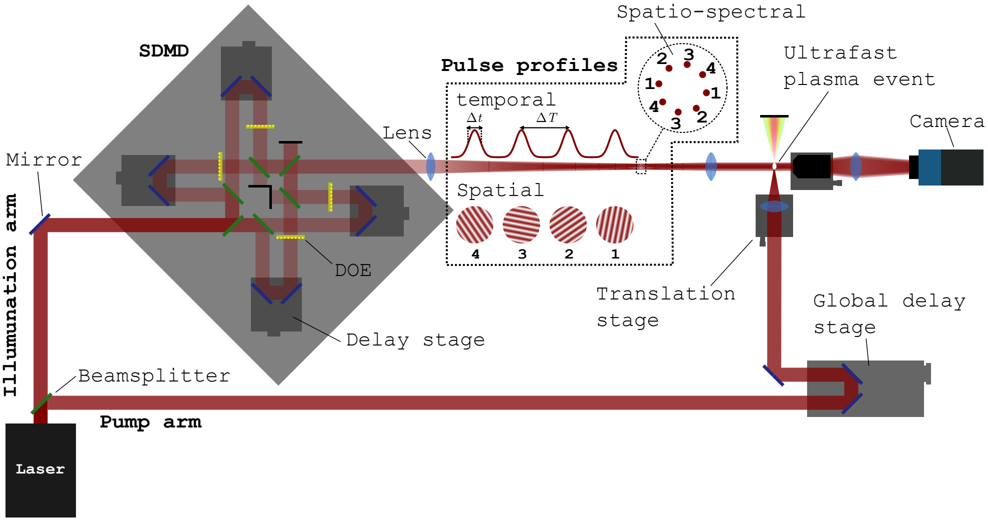


Figure 4.14: The optical setup used throught the ultrafast experiments in Paper V. An initial short laser pulse is split up into a pump arm and a illumination ar. The pump arm consists of a global delay stage and adjustable mirrors, to control the timing and position of the laser induced plasma event. The illumination arm consists of the split, delay and modulate device (SDMD) that shapes the temporal, spatial and spatio-sepctral profiles of the illumination pulses, as well as the camera and optics for imaging.

When integrated in the complete setup the SDMD enabled sufficiently high precision in both space and time. The data acquired using this device includes the 1 Tfps sequence shown in Fig. 4.15, as well as all imaging data presented in Paper v. In the sequence, a laser-induced signal in the form of refractive index variations can be seen to evolve and propagate at the speed of light over approximately 1 mm. Several distinct features are visible in the signal, including a central filament (arrow 1), the outer edge of the driving laser pulse (arrow 2), a contourless widening of the central filament near the geometrical focus (arrow 3), the appearance of multiple short filaments (arrow 4), and a single surviving weak filament (arrow 5). The combination of high spatial detail and ultrafast temporal resolution demonstrates the general capabilities of FRAME and the particular functionality of the 3D printed SDMD. However, aligning using hex screws was more difficult than necessary and is not something I recommend. Using rubber bands was, obviously, a major mistake. I have since bought steel springs but have not yet redesigned the mirror holders to integrate them. Whether or not the device will feature in future experiments remains to be seen.

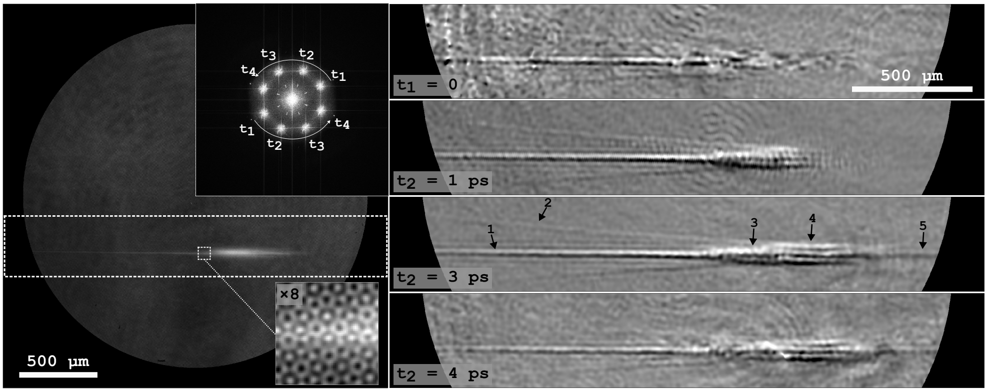


Figure 4.15: An image sequence of a laser-induced plasma event, including the formation and propagation of a central filament, acquired at 1 Tfps. The detector image and its Fourier transform are displayed on the left, while the reconstructed phase images are shown on the right.

3.2 Phase unwrapping

When analyzing FRAME data, the phase change at any pixel position $\vec{r}$ is found as the inverse tangent of a ratio between two numbers (see Eq. 4.3). In principle the true phase change $\Delta\phi(\vec{r})$ can take on any arbitrarily large positive or negative real value. However, a phase calculated as $\phi = \tan^{-1}(b/a)$, is always constrained to the interval $[-\pi, \pi]$. Consequently, the *measured* phase $\Delta\phi' = \phi - \phi_0$ will always fall within the interval $[-2\pi, 2\pi]$. If the true phase $\Delta\phi$ lies outside this range, it becomes *wrapped*, meaning that the measured phase $\Delta\phi'$ differs from the true phase $\Delta\phi$ by a multiple of 2π:

$$\Delta\phi' = \Delta\phi + m \cdot 2\pi, \tag{4.4}$$

where m is an integer. Hence, in regions where $\Delta\phi$ exceeds the interval $[-2\pi, 2\pi]$, wrapping introduces discontinuities in the measured data, as shown in Fig. 4.16a. The true phase change (blue curve) varies smoothly across the x-axis, while the measured phase (red curve) exhibits abrupt 2π jumps.

Fortunately, unwrapping the signal in such cases is straightforward. By iterating through the sequence and adjusting each value by the integer multiple of 2π that minimizes its difference from the previous value, one can restore the continuity of the phase. The result, shown as the unwrapped green dashed curve in Fig. 4.16a, retains the same shape as the true phase and differs only by a constant offset [136].

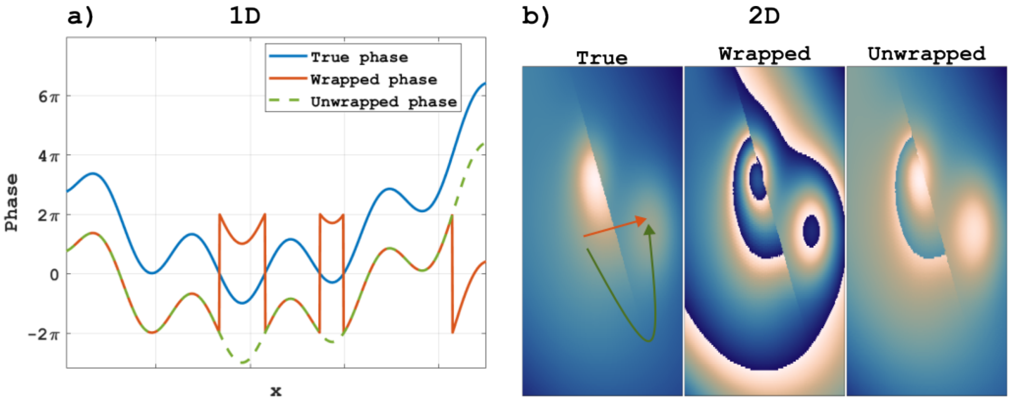


Figure 4.16: Wrapping of phase data in 1 and 2 dimensions. (a) True, wrapped and unwrapped 1D phase data. (b) True, wrapped and unwrapped 2D phase data with an unsolvable discontinuity.

3.2.1 The two dimensional case

Reconstructed FRAME signals are, however, inherently 2D, and in this case unwrapping is not always straightforward. In the central image in Fig. 4.16b wrapped synthetic phase data is presented. It is similar to what can be obtained experimentally using FRAME and requires unwrapping to be presentable. One [15] might naively think that the same strategy as in the 1D case is applicable: iteratively adjusting each pixel according to the value of its previous neighbor.

The problem with this approach is that each pixel in 2D has two previous neighbors (e.g., one below and one to the left), and these neighbors might not agree on the correction that should be applied. Another way to illustrate the issue is to consider that any two pixels are connected by more than one possible path. There might be one smooth path without discontinuities (the curved green arrow in Fig. 4.16b), but also another with a 2π jump (the straight red arrow). If the discontinuity in the second path is corrected, a new discontinuity is introduced in the first. This creates an inherent ambiguity: the problem cannot be fully resolved without compromise.

This challenge is why complex 2D unwrapping algorithms have been developed, designed to make these compromises and perform the task as efficiently as possible [137]. The algorithm used most extensively during this work was introduced under the title *Fast two-dimensional phase-unwrapping algorithm based on sorting by reliability following a noncontinuous path* in 2002 [138]. In this context it performed as well as the alternatives that were tested and was faster. Speed might not be a critical factor when analyzing results in the office, but during lab work, it can be essential to see signal changes in near real time during the alignment process. Despite being the fastest alternative, the algorithm was still too slow for this

---

[15] Old me

purpose.

3.2.2 A new 2D unwrapping strategy

To address the limitations of 2D unwrapping algorithms during lab work, I have developed a fast unwrapping strategy to apply existing algorithms to wrapped data more efficiently. This approach has sped up unwrapping by more than an order of magnitude in the present case, without noticeable differences in the results. The essence of the method lies in performing the computationally intensive unwrapping on a down-sampled version of the data to quickly approximate the unwrapped phase. This approximation is then up-sampled and utilized to find the correction needed to accurately unwrap the wrapped data, without loss of finer details.

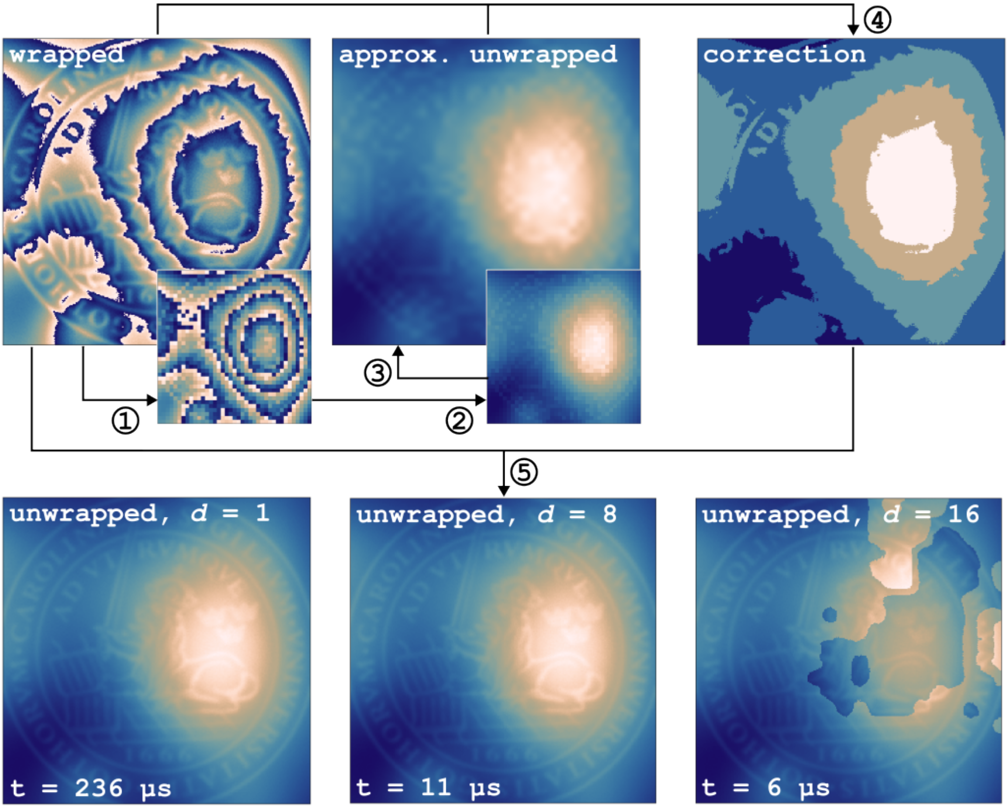


Figure 4.17: Overview of the fast unwrapping strategy. (1) The wrapped phase data is down-sampled by a factor d. (2) The coarse wrapped phase is unwrapped using a 2D unwrapping algorithm. (3) The unwrapped coarse phase is up-sampled to approximate the full-resolution unwrapped phase. (4) The correction is calculated as the difference between the wrapped and approximate unwrapped data and rounded to the nearest integer multiple of 2π. (5) The wrapped phase is unwrapped by adding the correction, resulting in the final high-resolution unwrapped phase. At the bottom are three results, obtained using different down-sampling factors, with indicated computational times.

The full strategy can be followed in Fig. 4.17, where wrapped synthetic data is unwrapped in five steps.

Step 1 is to downsample the image (data) by a user-defined downsampling factor d. It is important that d is set so that the rough wrapping pattern is still resolved. However, the finer details of the image do not need to be resolved. In the illustration of the involved steps in Fig. 4.17 $d = 8$, but the resulting unwrapped images for $d = 1$ and $d = 16$ are also shown for reference. When down-sampling, no interpolation must be performed; instead, pixel values should be set to the value of the nearest pixel in the full-resolution image. Any interpolation would smooth the sharp wrapping edges, preventing them from being correctly identified in the next step.

Step 2 is to unwrap the downsampled image using an existing 2D unwrapping algorithm. Any algorithm can be used, but its performance will influence both the quality of the final result and the computational time. In this step, it is crucial that the unwrapping is successful; otherwise, an unsatisfactory end result will follow.

Step 3 has the now unwrapped downsampled data upsampled back to full resolution. Here, interpolation is applied to create a smooth approximation of the unwrapped data. In the example shown, the result of this step contains no discontinuities and successfully captures the low-frequency components of the final unwrapped image. [16]

Step 4 involves calculating the correction using the initial wrapped image and the approximate unwrapped image. Specifically, the correction is calculated by subtracting the wrapped data from the approximate unwrapped data and rounding to the nearest integer multiple of 2π. As long as the approximate unwrapped data is within π of the true unwrapped data, this step will produce the correct correction. This follows directly from Eq. 4.4, where the correction must be an integer multiple of 2π.

Step 5 is the final step, in which the correction is added to the wrapped image to obtain the unwrapped image. If all steps were performed successfully, the resulting image will have no discontinuities while preserving the high-frequency components of the detector image. In the example shown, the unwrapped image obtained using $d = 8$ is identical to the image obtained without downsampling ($d = 1$), i.e. using the standard unwrapping procedure, despite requiring less than 5% of the computational time. The computational time can be further reduced by increasing the downsampling factor, but unwrapping will eventually fail when the wrapping pattern is no longer properly resolved.

The MATLAB code for each step is provided below.

---

[16] If you squint, the approximate unwrapped image is indistinguishable from the unwrapped image just below it in the figure. This indicates that they are identical in terms of low-frequency components.

```
wrapped_small = imresize(wrapped, 1/d, 'nearest');             % step 1
unwrapped_small = phase_unwrap(wrapped_small);                  % step 2
approxUnwrapped = imresize(unwrapped_small, d, 'bicubic');     % step 3
correction = round((approxUnwrapped-wrapped)/2/pi)*2*pi;       % step 4
unwrapped = wrapped + correction;                                % step 5
```

To evaluate how well this faster unwrapping strategy performs on experimental data, it was applied to 10 datasets, each consisting of four phase images. Each image was first unwrapped without down-sampling to establish a ground truth. To introduce more phase wraps, the resulting unwrapped image was multiplied by 5, and normally distributed noise (with a standard deviation of 0.05 cycles) was added before being wrapped. An example of one of the wrapped test images is shown in Fig. 4.18a.

Each test image was unwrapped with increasing down-sampling factors, ranging from 1 to 32. The computational time for each unwrapping was recorded, and the proportion of pixels deviating from the ground truth by more than π was counted as faulty. Examples of the resulting unwraps are shown in Fig. 4.18b–d for $d = 1$, $d = 8$, and $d = 16$, respectively. For $d = 1$, the unwrapping is perfectly successful. When $d = 8$, small regions in the lower-left part of the image become incorrectly unwrapped, and for $d = 16$ these incorrectly unwrapped areas grow even larger.

The results are summarized in Fig. 4.18e–f, where the effect of d on the proportion of erroneous pixels and the computational time is illustrated. For example as d increases from 1 to 4, the average computational time decreases from 5.1 to 0.38 s, while the proportion of incorrectly unwrapped pixels increases only slightly (from 0% to 0.03%). Thus, by employing this faster unwrapping strategy, computational time can be significantly reduced with only a minor loss in accuracy. While this trade-off may be unacceptable for certain applications, in others it may be well worth the gain in speed. From my own experience, as long as the majority of wrapping artifacts originate from the background rather than the actual signal, using this strategy with $d \approx 4$ produces no noticeable difference in the final result and significantly fewer computer-induced delays in the experimental work.

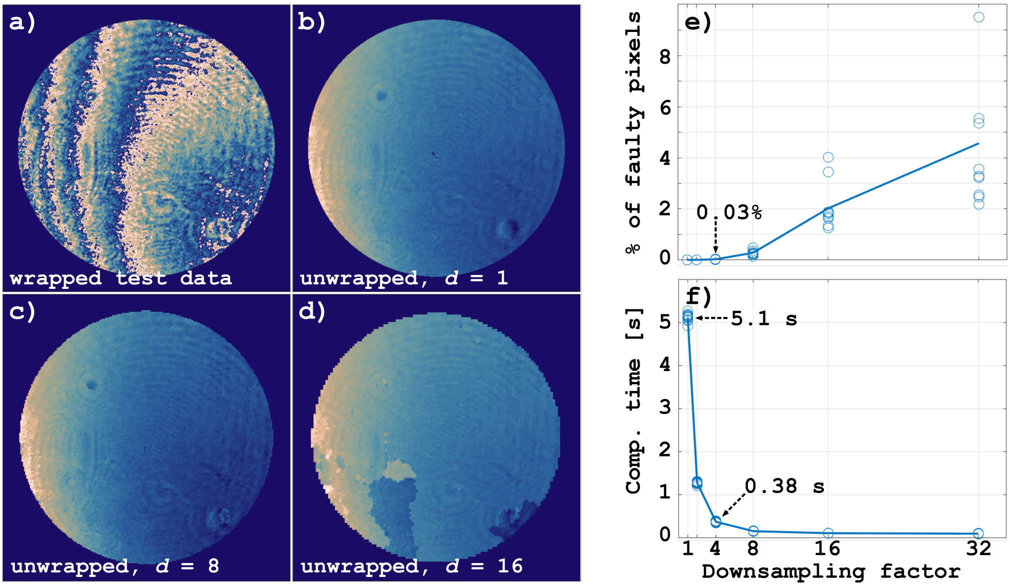


Figure 4.18: Quantitative effects of increasing the down-sampling factor on phase unwrapping. (a) An example of a wrapped test image after introducing additional phase wraps and noise. (b–d) Unwrapped results for down-sampling factors $d = 1, 8$, and 16, respectively, illustrating how incorrectly unwrapped regions appear and expand in the lower areas as d increases. (e) The percentage of pixels deviating from the ground truth by more than π as a function of d. (f) The computational time as a function of d.

3.3 Making sense of the data

When analyzing the data for Paper v, I realized that the phase signal is likely not only, or perhaps not even primarily, caused by light deflection, as in traditional schlieren imaging. Instead, the dominant effect is probably target-induced optical density variations causing light to be delayed. The primary mechanism is thus likely *interferometric*. In hindsight, this is a reasonable conclusion, given that the modulation patterns ultimately arise from *interference* when performing FRAME experiments with laser light.

The reason this realization came late is that the first phase-sensitive experiments (in Paper III) were conducted using incoherent illumination, eliminating all interferometric effects. Consequently, the observed phase changes in the pattern were solely due to deflection, as in conventional schlieren imaging. However, when using coherent light, the phase change appears identical, making it difficult to determine the underlying mechanism. To confirm that the observed effect is indeed primarily interferometric, I constructed a 3D ray-tracing model designed to replicate the experimental conditions in Paper v.

In the model, a pair of probe pulses, consisting of a user-defined number of evenly spaced rays, interact with an orthogonally propagating target pulse. The two probe pulses correspond to one temporal instance and result in one reconstructed frame. The situation is illustrated in Fig. 4.19a, with arrows indicating the direction of the pulses. Initially, the directions are identical for all rays within a sub-pulse and are defined such that the pulses will perfectly overlap at the plane $x = 0$.

The rays are propagated one time step at a time using a fourth-order Runge-Kutta method until they reach a user-defined output plane ($x = 0.5$ mm in this case). As the rays propagate, they experience an altered refractive index due to the high-intensity target pulse. The target pulse is modeled as a propagating and focusing 3D refractive index field, and at each time step, the field values are interpolated to the positions of the probing rays. The target pulse can thus be likened to a radially shrinking lens moving at the speed of light. From the 3D gradient of the refractive index field, the change in direction of each ray can be determined, and their positions and directions are subsequently updated. Ordinary ray-tracing programs typically assume stationary optical components, but for this model to be accurate, the light-bending optical component must propagate as fast as the probe pulses.

As the probing rays reach the output plane, their time spent in the interaction volume, their positions, and directions are recorded. From the positions and directions, their projected positions at a user-defined image plane are computed. At this plane, the two probe pulses are synthetically made to interfere, with the local phase of the interference pattern determined by the exit times of the rays. One such interference pattern is displayed in Fig. 4.19b, along with its Fourier transform. As a sanity check, it can be confirmed that the angle θ of the interference pattern is identical to the rotation angle of the pulses in the interaction volume in Fig. 4.19a, and that the period matches the value predicted by the convergence angle of the pulses.

Depending on the position of the image plane, the appearance of the reconstructed phase images will vary. With an image plane far from the target plane, the observed structure will be larger, and the dynamic range of the phase signal will be higher. This effect is quantified in Fig. 4.19c, where dynamic range and size are plotted against the image plane offset. These behaviours are also observed in experimental data, for example, in figure 6 of Paper v. The magnitude of the refractive index field has been adjusted so that the dynamic range of the model phase images approximately match what we observe under experimental conditions. [17]

In order to quantify how much of the observed phase is a result of deflection, the light bending properties of the model was turned off and the result was compared with previous results. As illustrated in Fig. 4.19d, the difference in observed phase when deflection of rays is accounted for and when it is not is on the order of 10^{-8} radians, which is negligible in comparison to the actual phase shift of approximately five radians. It can therefore be

---

[17] For the images in Fig. 4.19f the dynamic range is approximately 0.7. These images were acquired with an image plane offset near zero, and their dynamic range should hence be compared to the minimum modeled dynamic range in Fig. 4.19c, which is sligtly below 1.

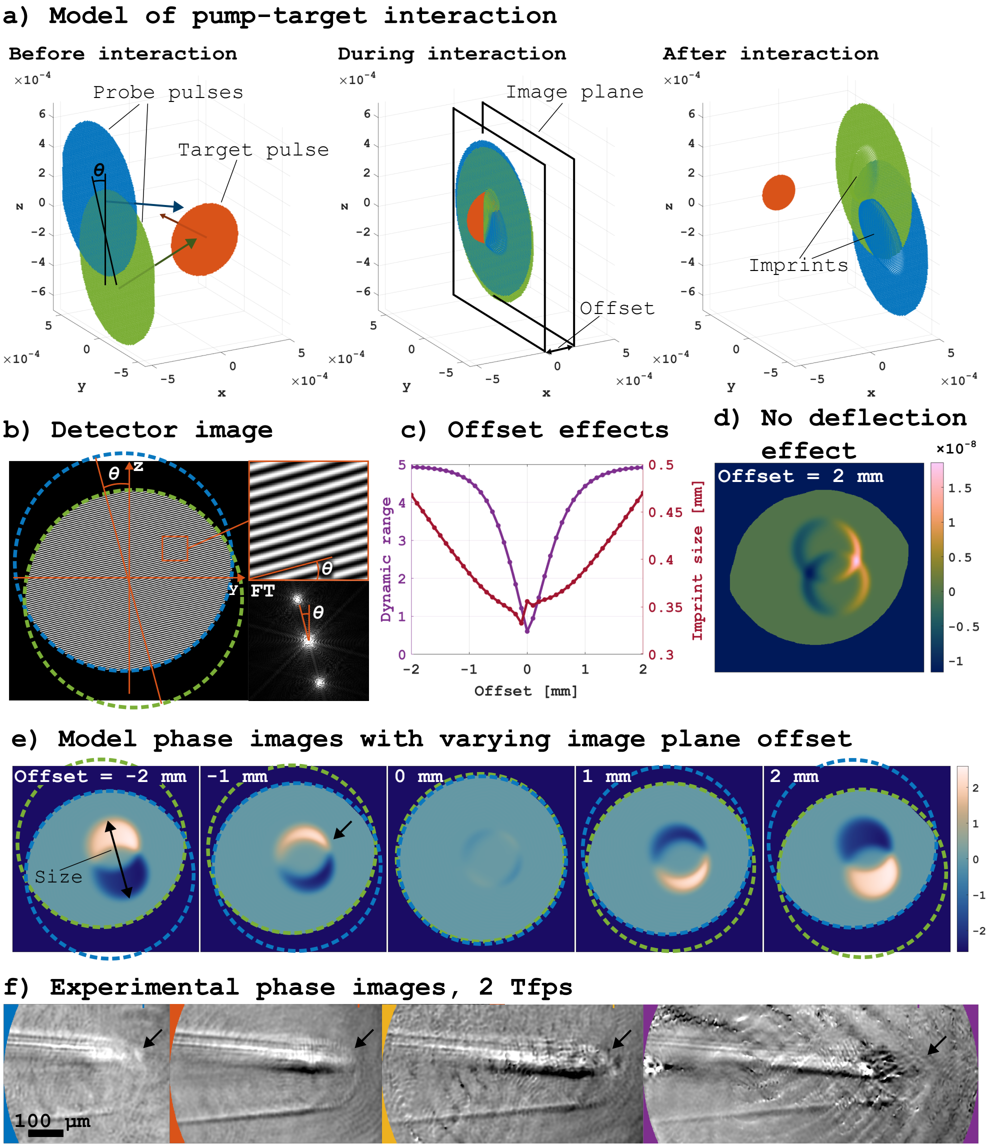


Figure 4.19: Model and analysis of the interaction between a focusing high-intensity target pulse and two converging probe sub-pulses, which subsequently interfere to produce an interference pattern. (a) Visualization of the interaction between the sub-pulses and the target pulse. The arrows indicate the propagation directions of each pulse. The interaction induces a delay in the probe pulses, leaving visible imprints. (b) The detector image, generated by projecting each sub-pulse onto a user-defined image plane after interacting with the target pulse and computing the resulting interference pattern. The rings indicate the positions of the pulses, which vary with the image plane offset. (c) The dynamic range and size of the structure observed in the reconstructed phase images in (e) as a function of the image plane offset. (d) Difference in detected phase change when deflection of rays is accounted for and when it is not. (e) Phase images reconstructed from synthetic detector images, similar to (b). The black arrow highlights a rounded edge in the direction of propagation, also observed in the experimental data in (e). (f) A 2 Tfps phase image sequence, showing the same rounded leading edge (black arrow) as seen in the modeled phase images.

concluded that the vast majority of the observed phase shift in the FRAME experiments in Paper v is due to delay (interferometric effects), rather than deflection of rays.

Another feature explained by the model is a finding made when examining phase images from Paper v, acquired under experimental conditions similar to those in the model. The observation was that the leading edge of the trace left by the propagating target pulse is curved, as marked by black arrows in the experimental phase images in Fig. 4.19f. This is somewhat peculiar, as the target pulse is so short in duration ($\sim$ 30 fs) that it is essentially flat ($\sim$ 10 µm), resembling a pancake. Seen from the side one could expect it to appear as a line. However, a similar rounded edge is also seen (and marked by a black arrow) in the model phase images in Fig. 4.19e. The reason for the curved imprint becomes apparent when a slow-motion simulation of the interaction between the target and probe pulses is examined. Due to the orthogonal—but otherwise identical—forward motion of the target and probe pulses, each vertical slice (in the x-z plane) of the target pulse interacts with a different vertical slice (in the y-z plane) of the probe pulses. As a result, it is in practice the front, rather than the side, of the target pulse that becomes imprinted onto the probe pulses.

It was actually this observation that motivated the development of the model, to determine the shape of the imprint in the probe pulses. The model images in Fig. 4.19e show the same curved leading edge but lack the trailing features seen in the experimental images. This is because, in the model, there are no lingering effects—the target pulse only affects the probe pulses when they overlap. In reality, the matter behind the target pulse is altered by it, leaving a trace behind the pulse. Even if this effect is not accounted for the model replicates the rounded leading edge. By allowing the dynamic interaction between probe pulses and target pulse to be visualized the reason for the rounded edge becomes very clear.

Chapter summary

In this chapter, we have investigated the achievable sequence depth of FRAME and demonstrated that it can extend to at least several hundred frames. Additionally, we have discovered that FRAME inherently possesses the ability to extract phase information, which we have applied to ultrafast imaging by visualizing the propagation of laser filaments in air. However, most of the results from the papers I have contributed to are not detailed here but can be found in the attached publications. Instead, this thesis has focused on the underlying work leading to those results and the rationale behind our experimental and methodological choices. In some cases, I have expanded on the discussion of key findings or provided an improved analysis. While I have placed particular emphasis on my own contributions, all of this work has been a team effort.

In the next and final chapter, I will place these developments into a broader context, considering their significance within the wider field of ultrafast imaging.

Chapter 5

Conclusion and outlook

1 FRAME and the history of imaging.

In the first chapter of this thesis, I broadly outlined the history of imaging, focusing on major milestones such as the advent of cameras and digital imaging. These milestones represent qualitative advancements in the sense that they introduced entirely new capabilities, which in some cases completely transformed the way we acquire image data. If such a breakthrough proves viable, it is typically followed by quantitative improvements over the years to come. These refinements do not introduce fundamentally new capabilities but incrementally enhance existing features such as pixel size, sensitivity and shutter speed, which in turn directly improve the resolution, contrast, and exposure time. It is crucial to recognize that without these painstaking, sometimes tedious, quantitative improvements, qualitative breakthroughs would amount to little, as most techniques are not born mature and ready for use.

When FRAME was first introduced the use of short laser pulses for optically multiplexing ultrafast image sequences was not an entirely new concept[100, 99], meaning that the fundamental milestone had already been achieved. As for breakthroughs in general, this milestone was not about a specific technique but rather about managing to decouple the frame rate of a video sequence from the camera speed, thus allowing frame rates to exceed those imposed by sensor readout speed. FRAME also achieved this decoupling, but through a new technical solution, building on the inventors previous work in structured illumination [139, 140, 141, 142]. While FRAME itself may not represent a major qualitative leap in ultrafast imaging capabilities, its specific technical implementation was qualitatively new. Consequently, FRAME should be viewed as part of the broader trend of employing ultrafast lasers for single-shot ultrafast imaging, and as such as a significant contribution to the ongoing quantitative refinement of these techniques.

This overarching development is unlikely to be a dead end, and therefore we should expect continuous improvements over time. Given that this era of imaging is only about a decade old, it is unreasonable to expect perfect image quality at this stage. Hence, when someone in the future writes the complete history of imaging, STAMP, CUP, FIP, and FRAME images from today will likely (and hopefully) be treated similarly to how the earliest daguerreotypes or the first digital images (see Fig. 1.2c and Fig. 1.3a) are viewed in hindsight, i.e. as a starting point for what would become an indispensable imaging technique.

This perspective should also be applied to the contributions of this thesis; as incremental but meaningful improvements to ultrafast imaging in general and FRAME in particular. One such improvement is the demonstration of much longer sequence depths. This result will likely serve as an encouragement, and insurance that it is technically feasible, to those that in the future will try to achieve long FRAME sequences in the more challenging ultrafast regime. The one improvement that could be considered qualitative, in the sense described above, is the introduction of phase sensitivity to FRAME. However, this is only a qualitative step forward within the narrow scope of FRAME itself, as other ultrafast techniques had already demonstrated phase sensitivity [143, 117]. That said, the level of detail we achieved in our ultrafast imaging of plasma events in Paper v represents an improvement over previous studies—perhaps not a qualitative leap in the ultrafast imaging community, but certainly a significant step forward, and as such, an important contribution.

2 FRAME and other ultrafast imaging techniques

One sign that we are indeed in the early stages of the ultrafast imaging era is the sheer number of techniques currently being developed. We can compare this situation to the early days of photography, when different photosensitive materials and formulas were being tested until a clearly superior solution emerged. We might be in a similar situation today, where the search is ongoing for an ultimate solution that will surpass all others and dominate ultrafast imaging for the foreseeable future. At present, no such solution appears to exist. To the best of my knowledge, no single technique has yet demonstrated ultrahigh temporal resolution while simultaneously achieving high spatial resolution, long sequence depth, large field of view, and strong sensitivity to relevant signals. Each existing method has inherent limitations. For instance, STAMP relies on probe pulses with different wavelengths, making it unsuitable for spectrally sensitive measurements. CUP suffers from a trade-off between spatio-temporal resolution and signal strength due to electron repulsion in the streak camera. FRAME, while versatile, cannot be operated passively on an ultrafast timescale, and is yet to demonstrate long sequences in the ultrafast regime. Given these constraints, it remains uncertain whether a single technique will ever dominate, or if trade-offs will always necessitate the approach to be tailored to the specific application.

Being in this new imaging regime means that the field is somewhat uncharted, and unfortunately, that it lacks a golden standard for defining what is good ultrafast imaging. This, and the realization that imaging speeds are no longer constrained by electronic limitations, is likely the reason that so much emphasis has been placed on achieving the highest possible frame rates. This focus is evident from numerous explicit claims of world record speeds, including from our own group [99, 130, 115], and from the frequent inclusion of frame rate values in paper titles [144, 88, 109, 108, 90]. While frame rate is undoubtedly an important parameter, it should not be the sole criterion for evaluating an ultrafast imaging technique. Doing so is somewhat analogous to stating the maximum RPM of a car; while relevant, the actual speed of the car is also influenced by multiple other factors.

From my own experience, and from looking at both my own and others' data, the parameter that is often lacking is actually not frame rate. I have rarely contemplated an ultrafast image sequence and thought: "If only the frame rate was higher!" Instead, I have wished for longer sequences and better image quality in the individual frames. This certainly applies to the ultrafast data presented in this thesis. In my view, the scientific value of the image sequences in, for example, Figures 4.15 and 4.19f would benefit far more from improved image quality than from a higher frame rate. An increased frame rate would probably be more of technical than practical significance.

To facilitate comparisons between different techniques, I have argued that the temporal dimension should be treated analogously to the spatial dimension, and that the most important temporal parameters are temporal resolution [1] and sequence depth, which correspond to spatial resolution and field of view, respectively, in the spatial domain. When evaluating which technique performs best, these parameters should take precedence over metrics such as frame rate and number of pixels. Ideally, a standardized benchmarking test should be established, in which spatial and temporal resolution are measured from a dynamic target. To enable measurements of the highest temporal resolutions, this target must be ultrafast—practically, it would have to be an instantaneous laser-induced event, such as filamentation.

If I were forced to make an argument for why FRAME might be the strongest candidate for ultrafast imaging, I would emphasize three key aspects. First, it is relatively easy to use. The only fundamental requirement is that the target must be illuminated with spatially modulated light, which is achieved simply by imaging a grating or a diffractive optical element (DOE) onto the target. This is significantly simpler than, otherwise similar, ultrafast holographic methods, where each probe pulse must be both spatially and temporally aligned to match a separate reference pulse for interference pattern generation. Also, the

---

[1] An ultrahigh temporal resolution inherently requires an ultrahigh frame rate, while ultrahigh frame rate does not necessarily imply high temporal resolution.

post-processing in FRAME is based on Fourier analysis, which is computationally efficient and can be implemented in fewer than ten lines of code. [2] This allows for nearly instant access to processed data during experiments, greatly enhancing workflow in the laboratory.

Second, FRAME is highly versatile. It can be implemented with a wide range of wavelengths, which can either be identical or unique across the probe pulses. In contrast, some other techniques require probe pulses to be spectrally distinct, which not only complicate spectrally sensitive measurements, but also require broadband light-sources. FRAME imposes no such constraint and can be applied across vastly different timescales, as well as in both microscopic and macroscopic configurations, as demonstrated in this work. An additional advantage of FRAME's compatibility with various light sources and pulse durations is that new operators can gain confidence and understanding using more affordable and less hazardous lasers before committing to ultrafast operation.

Third, FRAME is inherently highly sensitive to refractive index gradients and has been demonstrated with high spatial and temporal resolution in an ultrafast configuration. These properties have been specifically demonstrated in the context of this work and are of particular importance for ultrafast imaging, as potential ultrafast targets tend to be transparent and microscopic in scale.

3 FRAME and the Future

As already emphasized, it is still too early to determine which technique—if any—will become dominant in the field of ultrafast imaging. Ultimately, the defining factor for a technique's success will not be a performance statistic, but its scientific impact in fields beyond ultrafast imaging. In other words, for FRAME to remain relevant, it is not enough to demonstrate ever-improving performance if that performance is not put to practical use. Ideally, we want researchers from other fields to adopt the technique to address existing scientific questions. Naturally, the technique itself must undergo technical development before widespread adoption can occur, but at some point, a transition from technical research to applied research must take place. Hopefully, we are at the beginning of this transition, and in the coming years, we will see more research groups utilizing FRAME as a diagnostic tool for their specific studies. In the long run, such a transition will likely accelerate the technical development of FRAME, as broader adoption will lead to increased investment in the technique.

To facilitate the transition from technical to applied research, two major developments are necessary. The first concerns the quality of the data produced in FRAME experiments. While improvements are needed, the focus should not be on increasing the frame rate

---

[2] In Matlab, Python or similar high-level language.

but rather on enhancing the image quality of individual frames. This includes improving contrast and spatial resolution to ensure that the finest details of dynamic targets can be observed and accurately traced. To minimize motion blur, improving temporal resolution may also be necessary, but the primary goal should be to make the reconstructed frames as sharp as possible.

The second consideration is ease of operation. When designing the split, delay, and modulate device (SDMD) for the setup in Paper v, the goal was to create a ”FRAME box” [3] —a device that could be easily implemented in various experimental setups. I believe this remains a valid goal and should serve as a guiding principle in the coming years. If FRAME is to be widely adopted, it must be easy to set up and operate. However, I believe there is significant room for improvement beyond the SDMD developed in this work. For example, by sacrificing some temporal control, it should be possible to drastically reduce the number of components, making the system much easier to use.

Ideally, the device should be a closed unit, perceived by the user as a single component—much like a camera lens. This ”FRAME tube” could be designed so that a single input pulse exits as ten (or so) temporally distinct probe pulses with unique modulations. A complete experimental setup would then ”only” require an ultrashort laser source, a few mirrors for alignment, the FRAME tube, and a camera. One possible approach to constructing the FRAME tube would involve a diffractive optical element (DOE), followed by a stepped glass sheet, and a lens. The DOE would split the input pulse into sub-pulses at different angles (defining the spatial frequencies), the stepped glass sheet would delay the sub-pulses pairwise (setting the timing), and the lens would project the sub-pulses onto the target (effectively imaging the DOE onto the target). Such a device would likely make FRAME an even more compelling diagnostic tool and further its impact across relevant scientific fields.

All in all, this aligns with the ultimate goal of our research: to contribute to the ongoing progression that gained momentum during the scientific revolution, when the importance of observation was realized, and tools such as microscopes and telescopes enabled deeper insights into the natural world. Whether FRAME will become such an iconic instrument in the hands of researchers remains an open question. However, the pursuit of expanding our observational capabilities has been the fundamental driving force behind the research presented in this thesis, and seeing FRAME contribute to future scientific breakthroughs would be the ultimate reward.

---

[3] That name, though descriptive, never stuck with me.

Chapter 6

Appendix

Appendix A: Resolution from edge

When analyzing how a perfect step edge (from intensity 0 to 1) is blurred by a Gaussian point-spread function of standard deviation σ, we find that the resulting intensity profile can be written as

$$E(x) = \tfrac{1}{2}\left[\, 1 + \mathrm{erf}\big(\tfrac{x}{\sqrt{2}\,\sigma}\big)\right],$$

where, erf is the error function, which arises naturally from integrating the Gaussian function e^{-t^2} and is defined by

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, \mathrm{d}t.$$

From the blurred edge expression, the points at which the intensity is 75% and 25% of the maximum occur at

$$x_{75\%} \approx +0.675\,\sigma \quad \text{and} \quad x_{25\%} \approx -0.675\,\sigma,$$

so the distance over which the intensity drops from 75% to 25% is

$$d_{50\%} = x_{75\%} - x_{25\%} = 1.35\,\sigma.$$

A Gaussian also has a known relationship between its standard deviation and its full width at half maximum (FWHM):

$$\mathrm{FWHM} = 2\sqrt{2\,\ln(2)}\,\sigma \approx 2.355\,\sigma.$$

According to Lord Rayleigh's criteria two identical Gaussian spots are just resolved when they are separated by $1.19 \times \mathrm{FWHM}$ (see Equations 2.1 and 2.2). Calling this minimal resolvable separation r, we therefore have

$$r = 1.19 \cdot 2.355\,\sigma = 2.80\,\sigma,$$

and, since $d_{50\%} = 1.35\,\sigma$, it follows that

$$r = \frac{2.80}{1.35}\,d_{50\%} \approx 2.1\,d_{50\%}.$$

References

[1] Max Roser, Esteban Ortiz-Ospina, and Hannah Ritchie. Child mortality, 2023. Available at: `https://ourworldindata.org/child-mortality` (Accessed: December 10, 2024).

[2] Steven Shapin. *The scientific revolution*. University of Chicago press, 2018.

[3] Albert Einstein. *Essays in science*. Open Road Media, 2011.

[4] Aristotle. *Metaphysics*, volume 1. Harvard University Press, Cambridge, MA, 1933.

[5] Thomas Reid. *Essays on the intellectual powers of man*. J. Bartlett, 1850.

[6] René Descartes. *Meditations, objections, and replies*. Hackett Publishing, 2006.

[7] David Hume, editor. *An Enquiry Concerning Human Understanding and Other Writings*. Cambridge University Press, New York, 2007.

[8] John Locke. An essay concerning human understanding, 1690. In W. Dennis, editor, *Readings in the History of Psychology*, pages 55–68. Appleton-Century-Crofts, 1948. Originally published in 1690.

[9] Kant Immanuel et al. Critique of pure reason. 1855.

[10] Fred Attneave. Some informational aspects of visual perception. *Psychological review*, 61(3):183, 1954.

[11] Richard L. Gregory. *Eye and Brain: The Psychology of Seeing*. Princeton University Press, 2015.

[12] David Prutchi. *Exploring Ultraviolet Photography: Bee Vision, Forensic Imaging, and Other NearUltraviolet Adventures with Your DSLR*. Amherst Media, 2016.

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

[14] Anne Sakdinawat and David Attwood. Nanoscale x-ray imaging. *Nature photonics*, 4(12):840–848, 2010.

[15] Martin Berger, Qiao Yang, and Andreas Maier. X-ray imaging. *Medical Imaging Systems: An Introductory Guide*, pages 119–145, 2018.

[16] Klaus M Pontoppidan, Jaclyn Barrientes, Claire Blome, Hannah Braun, Matthew Brown, Margaret Carruthers, Dan Coe, Joseph DePasquale, Néstor Espinoza, Macarena Garcia Marin, et al. The jwst early release observations. *The Astrophysical Journal Letters*, 936(1):L14, 2022.

[17] Bradley W Carroll and Dale A Ostlie. *An introduction to modern astrophysics*. Cambridge University Press, 2017.

[18] Yaron M Sigal, Ruobo Zhou, and Xiaowei Zhuang. Visualizing and discovering cellular structures with super-resolution microscopy. *Science*, 361(6405):880–887, 2018.

[19] Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Molecular Biology of the Cell*. Garland Science, 6th edition, 2014.

[20] James B Campbell and Randolph H Wynne. *Introduction to remote sensing*. Guilford press, 2011.

[21] Anthony Trewavas. The foundations of plant intelligence. *Interface focus*, 7(3):20160098, 2017.

[22] BBC. Plants fight one another, plants strangle one another!, January 2022. Accessed: 20-Mar-2025.

[23] Volker Sick. High speed imaging in fundamental and applied combustion research. *Proceedings of the Combustion Institute*, 34(2):3509–3530, 2013.

[24] Sigurdur T Thoroddsen, Takeharu Goji Etoh, and Kohsei Takehara. High-speed imaging of drops and bubbles. *Annu. Rev. Fluid Mech.*, 40(1):257–285, 2008.

[25] Michel Versluis. High-speed imaging in fluids. *Experiments in fluids*, 54:1–35, 2013.

[26] Dirk L Hoffmann, Christopher D Standish, Marcos García-Diez, Paul B Pettitt, James A Milton, João Zilhão, José Javier Alcolea-González, Pedro Cantalejo-Duarte, Hipólito Collado, Rodrigo De Balbín, et al. U-th dating of carbonate crusts reveals neandertal origin of iberian cave art. *Science*, 359(6378):912–915, 2018.

[27] Adam Brumm, Adhi Agus Oktaviana, Basran Burhan, Budianto Hakim, Rustan Lebe, Jian-xin Zhao, Priyatno Hadi Sulistyarto, Marlon Ririmasse, Shinatria Adhityatama, Iwan Sumantri, et al. Oldest cave art found in sulawesi. *Science Advances*, 7(3):eabd4648, 2021.

[28] Dick Harrison and Bo Eriksson. *Sveriges historia 1350–1600*. Norstedts Förlag, 2010.

[29] Thomas A Stoffregen. On the physical origins of inverted optic images. *Ecological Psychology*, 25(4):369–382, 2013.

[30] Filippo De Tomasi. Notes on camera obscura: three contemporary artistic perspectives on the path of photography. *Arte, Individuo y Sociedad*, 35(4), 2023.

[31] Mary Bellis. History of photography and the camera. *Retrieved April*, 11:2005, 2005.

[32] M Susan Barger and William B White. *The daguerreotype: Nineteenth-century technology and modern science*. JHU Press, 2000.

[33] James R Janesick, Tom Elliott, Stewart Collins, Morley M Blouke, and Jack Freeman. Scientific charge-coupled devices. *Optical Engineering*, 26(8):692–714, 1987.

[34] Fujifilm. Achievements - world's first digital camera, 2020. Available at: `https://web.archive.org/web/20201027124138/https://www.fujifilm.com/innovation/achievements/ds-1p/` (Accessed: December 10, 2024).

[35] National Institute of Standards and Technology (NIST). First digital image, 2022. Available at: `https://www.nist.gov/mathematics-statistics/first-digital-image` (Accessed: January 6, 2025).

[36] L. G. Knutsson. Sätt att för arbetsterapi e.d. framställa permanenta, mönstrade skivor samt mönstrad skiva framställd enligt sättet. Technical Report 217875, Swedish Patent Office, 1967. Swedish Patent 217 875. Filed April 24, 1962. Approved July 6, 1967. Published December 19, 1967. Available at: `https://upload.wikimedia.org/wikipedia/commons/f/f1/Swedish_patent_217875_S%C3%A4tt_att_f%C3%B6r_arbetsterapi.pdf`. Additional details at: `https://tc.prv.se/spd/patent?p2=3_IOQXsUr7M&content=217875&lang=sv&tab=1&hits=true&hitsstart=0&start=0`.

[37] Wikipedia contributors. Digital photography, 2025. Accessed: 2025-01-06.

[38] Michael Lesser. A summary of charge-coupled devices for astronomy. *Publications of the Astronomical Society of the Pacific*, 127(957):1097–1104, 2015.

[39] Ian S. McLean. *Electronic Imaging in Astronomy: Detectors and Instrumentation*. Springer, Dordrecht, Netherlands, 2nd edition, 2008.

[40] Abbas El Gamal and Helmy Eltoukhy. Cmos image sensors. *IEEE Circuits and Devices Magazine*, 21(3):6–20, 2005.

[41] MR Aziz, W Kuntjoro, et al. Terminal ballistic of aluminium plate: An experiment and numerical simulation. In *AIP Conference Proceedings*, volume 1774. AIP Publishing, 2016.

[42] Walter F Kosonocky, Guang Yang, Chao Ye, RK Kabra, Liansheng Xie, JL Lawrence, V Mastrocolla, FV 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. IEEE, 1996.

[43] TG Etoh, VTS Dao, HD Nguyen, K Fife, M Kureta, M Segawa, M Arai, and T Shinohara. R57 progress of ultra-high-speed image sensors with in-situ ccd storage. In *International image sensor workshop*. Intl. Image Sensor Society, 2011.

[44] 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, 2020.

[45] 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. *Physical review letters*, 119(21):214502, 2017.

[46] T Goji Etoh, Kazuhiro Shimonomura, Anh Quang Nguyen, Kosei Takehara, Yoshinari Kamakura, Paul Goetschalckx, Luc Haspeslagh, Piet De Moor, Vu Truong Son Dao, Hoang Dung Nguyen, et al. A 100 mfps image sensor for biological applications. In *High-Speed Biomedical Imaging and Spectroscopy III: Toward Big Data Instrumentation and Management*, volume 10505, pages 9–18. SPIE, 2018.

[47] 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. *Optics express*, 30(21):37664–37674, 2022.

[48] 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):483, 2017.

[49] Benjamin P Abbott, Richard Abbott, TDe Abbott, MR Abernathy, Fausto Acernese, Kendall Ackley, Carl Adams, Thomas Adams, Paolo Addesso, Rana X Adhikari, et al. Observation of gravitational waves from a binary black hole merger. *Physical review letters*, 116(6):061102, 2016.

[50] Michel Mayor and Didier Queloz. A jupiter-mass companion to a solar-type star. *nature*, 378(6555):355–359, 1995.

[51] Joseph John Thomson. Xl. cathode rays. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 44(269):293–316, 1897.

[52] James D Watson and Francis HC Crick. The structure of dna. In *Cold Spring Harbor symposia on quantitative biology*, volume 18, pages 123–131. Cold Spring Harbor Laboratory Press, 1953.

[53] Arno A Penzias and Robert W Wilson. A measurement of excess antenna temperature at 4080 mhz. In *A Source Book in Astronomy and Astrophysics, 1900–1975*, pages 873–876. Harvard University Press, 1979.

[54] Antony Hewish, S Joceyln Bell, John DH Pilkington, Paul Frederick Scott, and Robin Ashley Collins. Observation of a rapidly pulsating radio source. In *A Source Book in Astronomy and Astrophysics, 1900–1975*, pages 498–504. Harvard University Press, 1979.

[55] Dan Shechtman, Ilan Blech, Denis Gratias, and John W Cahn. Metallic phase with long-range orientational order and no translational symmetry. *Physical review letters*, 53(20):1951, 1984.

[56] JR Porter. Antony van leeuwenhoek: tercentenary of his discovery of bacteria. *Bacteriological reviews*, 40(2):260–269, 1976.

[57] Toshiki Tajima and John M Dawson. Laser electron accelerator. *Physical review letters*, 43(4):267, 1979.

[58] Nicholas H Matlis, S Reed, Stepan S Bulanov, Vladimir Chvykov, Galina Kalintchenko, T Matsuoka, P Rousseau, V Yanovsky, A Maksimchuk, S Kalmykov, et al. Snapshots of laser wakefields. *Nature Physics*, 2(11):749–753, 2006.

[59] Driss Oumbarek Espinos, Alexandre Rondepierre, Alexei Zhidkov, Naveen Pathak, Zhan Jin, Kai Huang, Nobuhiko Nakanii, Izuru Daito, Masaki Kando, and Tomonao Hosokai. Notable improvements on lwfa through precise laser wavefront tuning. *Scientific Reports*, 13(1):18466, 2023.

[60] Alexander Buck, Maria Nicolai, Karl Schmid, Chris MS Sears, Alexander Sävert, Julia M Mikhailova, Ferenc Krausz, Malte C Kaluza, and Laszlo Veisz. Real-time observation of laser-driven electron acceleration. *Nature Physics*, 7(7):543–548, 2011.

[61] Tianyang Yan and Lingfei Ji. Ultrafast laser filamentation in transparent solids. *Ultrafast Science*, 3:0023, 2023.

[62] A Sävert, SPD Mangles, M Schnell, E Siminos, Jason M Cole, M Leier, M Reuter, Matthew B Schwab, M Möller, K Poder, et al. Direct observation of the injection dynamics of a laser wakefield accelerator using few-femtosecond shadowgraphy. *Physical review letters*, 115(5):055002, 2015.

[63] A Braun, G Korn, X Liu, D Du, J Squier, and G Mourou. Self-channeling of high-peak-power femtosecond laser pulses in air. *Optics letters*, 20(1):73–75, 1995.

[64] Magali Durand, Aurélien Houard, Bernard Prade, André Mysyrowicz, Anne Durécu, Bernard Moreau, Didier Fleury, Olivier Vasseur, Hartmut Borchert, Karsten Diener, et al. Kilometer range filamentation. *Optics express*, 21(22):26836–26845, 2013.

[65] Jérôme Kasparian, Miguel Rodríguez, Guillaume Méjean, Jin Yu, Estelle Salmon, H Wille, R Bourayou, S Frey, Y-B André, André Mysyrowicz, et al. White-light filaments for atmospheric analysis. *Science*, 301(5629):61–64, 2003.

[66] Aurélien Houard, Pierre Walch, Thomas Produit, Victor Moreno, Benoit Mahieu, Antonio Sunjerga, Clemens Herkommer, Amirhossein Mostajabi, Ugo Andral, Yves-Bernard André, et al. Laser-guided lightning. *Nature photonics*, 17(3):231–235, 2023.

[67] Guillaume Schimmel, Thomas Produit, Denis Mongin, Jérôme Kasparian, and Jean-Pierre Wolf. Free space laser telecommunication through fog. *Optica*, 5(10):1338–1341, 2018.

[68] Jérôme Kasparian and Jean-Pierre Wolf. Physics and applications of atmospheric nonlinear optics and filamentation. *Optics express*, 16(1):466–493, 2008.

[69] CM Huntington, AGR Thomas, C McGuffey, T Matsuoka, V Chvykov, G Kalintchenko, S Kneip, Z Najmudin, C Palmer, V Yanovsky, et al. Current filamentation instability in laser wakefield accelerators. *Physical review letters*, 106(10):105001, 2011.

[70] P Béjot, Jérôme Kasparian, Stefano Henin, V Loriot, T Vieillard, E Hertz, O Faucher, B Lavorel, and J-P Wolf. Higher-order kerr terms allow ionization-free filamentation in gases. *Physical review letters*, 104(10):103903, 2010.

[71] Pengfei Qi, Lie Lin, Qiang Su, Nan Zhang, Lu Sun, and Weiwei Liu. In-situ visualization of multiple filament competition dynamic during nonlinear propagation of femtosecond laser. *Scientific Reports*, 7(1):10384, 2017.

[72] Peter Dombi and Martin Schultze. The nobel prize in physics 2023. *Europhysics News*, 54(5):8–9, 2023.

[73] S Bengtsson, ER Simpson, N Ibrakovic, S Ek, A Olofsson, T Causer, and J Mauritsson. Experimental observation of longer trajectories than previously observed in high-order harmonic generation. *Physical Review A*, 108(1):013505, 2023.

[74] Marcus Ossiander, Florian Siegrist, Vage Shirvanyan, R Pazourek, Annkatrin Sommer, Tobias Latka, Alexander Guggenmos, S Nagele, J Feist, J Burgdörfer, et al. Attosecond correlation dynamics. *Nature Physics*, 13(3):280–285, 2017.

[75] Marcus Ossiander, Johann Riemensberger, S Neppl, M Mittermair, Martin Schäffer, A Duensing, MS Wagner, R Heider, M Wurzer, M Gerl, et al. Absolute timing of the photoelectric effect. *Nature*, 561(7723):374–377, 2018.

[76] Rick Trebino and Daniel J Kane. Using phase retrieval to measure the intensity and phase of ultrashort pulses: frequency-resolved optical gating. *JOSA A*, 10(5):1101–1111, 1993.

[77] Chris Iaconis and Ian A Walmsley. Self-referencing spectral interferometry for measuring ultrashort optical pulses. *IEEE Journal of quantum electronics*, 35(4):501–509, 1999.

[78] Pierre-Marie Paul, Elena S Toma, Pierre Breger, Genevive Mullot, Frédérika Augé, Ph Balcou, Harm Geert Muller, and Pierre Agostini. Observation of a train of attosecond pulses from high harmonic generation. *Science*, 292(5522):1689–1692, 2001.

[79] Matthew S Robinson and Jochen Küpper. Unraveling the ultrafast dynamics of thermal-energy chemical reactions. *Physical Chemistry Chemical Physics*, 26(3):1587–1601, 2024.

[80] Hyotcherl Ihee, Vladimir A Lobastov, Udo M Gomez, Boyd M Goodson, Ramesh Srinivasan, Chong-Yu Ruan, and Ahmed H Zewail. Direct imaging of transient molecular structures with ultrafast diffraction. *Science*, 291(5503):458–462, 2001.

[81] 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, 2017.

[82] Rayleigh. Xxxi. investigations in optics, with special reference to the spectroscope. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 8(49):261–274, 1879.

[83] Sisi Zhou and Liang Jiang. Modern description of rayleigh's criterion. *Physical Review A*, 99(1):013808, 2019.

[84] Richard W Cole, Tushare Jinadasa, and Claire M Brown. Measuring and interpreting point spread functions to determine confocal microscope resolution and ensure quality control. *Nature protocols*, 6(12):1929–1941, 2011.

[85] Jinyang Liang, Peng Wang, Liren Zhu, and Lihong V Wang. Single-shot stereo-polarimetric compressed ultrafast photography for light-speed observation of high-dimensional optical transients with picosecond resolution. *Nature communications*, 11(1):5252, 2020.

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

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

[88] 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.

[89] Yongle Zhu, Xuanke Zeng, Yi Cai, Xiaowei Lu, Qifan Zhu, Liangwei Zeng, Tingchao He, Junzi Li, Yang Yang, Maijie Zheng, et al. All-optical high spatial-temporal resolution photography with raster principle at 2 trillion frames per second. *Optics Express*, 29(17):27298–27308, 2021.

[90] Liang Gao, Jinyang Liang, Chiye Li, and Lihong V Wang. Single-shot compressed ultrafast photography at one hundred billion frames per second. *Nature*, 516(7529):74–77, 2014.

[91] Chengshuai Yang, Dalong Qi, Jinyang Liang, Xing Wang, Fengyan Cao, Yilin He, Xiaoping Ouyang, Baoqiang Zhu, Wenlong Wen, Tianqing Jia, et al. Compressed ultrafast photography by multi-encoding imaging. *Laser Physics Letters*, 15(11):116202, 2018.

[92] Gary S Settles. *Schlieren and shadowgraph techniques: visualizing phenomena in transparent media*. Springer Science & Business Media, 2001.

[93] Stuart B Dalziel, Graham O Hughes, and Bruce R Sutherland. Synthetic schlieren. In *Proceedings of the 8th International Symposium on Flow Visualization*, volume 62. paper, 1998.

[94] Bruce R Sutherland, Stuart B Dalziel, Graham O Hughes, and PF Linden. Visualization and measurement of internal waves by 'synthetic schlieren'. part 1. vertically oscillating cylinder. *Journal of fluid mechanics*, 390:93–126, 1999.

[95] Peng Fei Gao, Gang Lei, and Cheng Zhi Huang. Dark-field microscopy: recent advances in accurate analysis and emerging applications. *Analytical Chemistry*, 93(11):4707–4726, 2021.

[96] Andreas Quirrenbach. Optical interferometry. *Annual Review of Astronomy and Astrophysics*, 39(1):353–401, 2001.

[97] Nils Abramson. Light-in-flight recording by holography. *Optics letters*, 3(4):121–123, 1978.

[98] Robert Curley et al. *Breakthroughs in Telephone Technology: From Bell to Smartphones*. Britannica Educational Publishing, 2011.

[99] Keiichi Nakagawa, Atsushi Iwasaki, Yu Oishi, Ryoichi Horisaki, Akira Tsukamoto, Aoi Nakamura, Kenichi Hirosawa, Hongen Liao, Takashi Ushida, Keisuke Goda, et al. Sequentially timed all-optical mapping photography (stamp). *Nature Photonics*, 8(9):695–700, 2014.

[100] Xiaolei Wang, Hongchen Zhai, and Guoguang Mu. Pulsed digital holography system recording ultrafast process of the femtosecond order. *Optics letters*, 31(11):1636–1638, 2006.

[101] 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. *Optics express*, 23(23):30512–30522, 2015.

[102] 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(1):66, 2022.

[103] Zhongxing Li, Long Xiao, Zonghao Feng, Zhiyuan Liu, Du Wang, and Cheng Lei. Sequentially timed all-optical mapping photography with quantitative phase imaging capability. *Optics Letters*, 49(18):5059–5062, 2024.

[104] Takakazu Suzuki, Ryohei Hida, Ryuta Ueda, Fumihiro Isa, Keiichi Nakagawa, and Fumihiko Kannari. Single-shot ultrafast 2d-burst imaging by stamp utilizing spectral filtering (sf-stamp). In *International Conference on Ultrafast Phenomena*, pages UTh4A–18. Optica Publishing Group, 2016.

[105] Takao Saiki, Takuya Hosobata, Yukihiro Kono, Masahiro Takeda, Ayumu Ishijima, Miu Tamamitsu, Yutaro Kitagawa, Keisuke Goda, Shin-ya Morita, Shinobu Ozaki, et al. Sequentially timed all-optical mapping photography boosted by a branched 4f system with a slicing mirror. *Optics Express*, 28(21):31914–31922, 2020.

[106] Takao Saiki, Keitaro Shimada, Ichiro Sakuma, Yuki Inada, and Keiichi Nakagawa. Single-shot ultrafast dual-view imaging of shock waves in parallel laser processing. *Applied Physics Express*, 16(9):092004, 2023.

[107] Xuanke Zeng, Shuiqin Zheng, Yi Cai, Hongyu Wang, Xiaowei Lu, Honggeng Wang, Jingzhen Li, Weixin Xie, and Shixiang Xu. Generation and imaging of a tunable ultrafast intensity-rotating optical field with a cycle down to femtosecond region. *High Power Laser Science and Engineering*, 8:e3, 2020.

[108] Qifan Zhu, Yi Cai, Xuanke Zeng, Hu Long, Liangwei Zeng, Yongle Zhu, Xiaowei Lu, and Jingzhen Li. Single-shot framing integration photography with high spatial resolution at 5.3* 1012 frames per second by an inversed 4f system. *arXiv preprint arXiv:2110.01941*, 2021.

[109] Qifan Zhu, Yi Cai, Xuanke Zeng, Hu Long, Hongyi Chen, Liangwei Zeng, Yongle Zhu, Xiaowei Lu, and Jingzhen Li. Fisi: frequency domain integration sequential imaging at 1.26× 10 13 frames per second and 108 lines per millimeter. *Optics Express*, 30(15):27429–27438, 2022.

[110] Yunhua Yao, Yilin He, Dalong Qi, Fengyan Cao, Jiali Yao, Pengpeng Ding, Chengzhi Jin, Xianyu Wu, Lianzhong Deng, Tianqing Jia, et al. Single-shot real-time ultrafast imaging of femtosecond laser fabrication. *Acs Photonics*, 8(3):738–744, 2021.

[111] Jiali Yao, Dalong Qi, Hongtao Liang, Yilin He, Yunhua Yao, Tianqing Jia, Yang Yang, Zhenrong Sun, and Shian Zhang. Exploring femtosecond laser ablation by snapshot ultrafast imaging and molecular dynamics simulation. *Ultrafast Science*, 2022.

[112] Emmanuel J Candes, Justin K Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 59(8):1207–1223, 2006.

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

[114] Liren Zhu, Yujia Chen, Jinyang Liang, Qiaofeng Xu, Liang Gao, Cheng Ma, and Lihong V Wang. Space-and intensity-constrained reconstruction for compressed ultrafast photography. *Optica*, 3(7):694–697, 2016.

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

[116] Taewoo Kim, Jinyang Liang, Liren Zhu, and Lihong V Wang. Picosecond-resolution phase-sensitive imaging of transparent objects in a single shot. *Science Advances*, 6(3):eaay6200, 2020.

[117] Yide Zhang, Binglin Shen, Tong Wu, Jerry Zhao, Joseph C Jing, Peng Wang, Kanomi Sasaki-Capela, William G Dunphy, David Garrett, Konstantin Maslov, et al. Ultrafast and hypersensitive phase imaging of propagating internodal current flows in myelinated axons and electromagnetic pulses in dielectrics. *Nature communications*, 13(1):5247, 2022.

[118] Peng Wang and Lihong V Wang. Single-shot reconfigurable femtosecond imaging of ultrafast optical dynamics. *Advanced Science*, 10(13):2207222, 2023.

[119] Yingming Lai, Yujia Xue, Christian-Yves Côté, Xianglei Liu, Antoine Laramée, Nicolas Jaouen, François Légaré, Lei Tian, and Jinyang Liang. Single-shot ultraviolet compressed ultrafast photography. *Laser & Photonics Reviews*, 14(10):2000122, 2020.

[120] Yu Lu, Terence TW Wong, Feng Chen, and Lidai Wang. Compressed ultrafast spectral-temporal photography. *Physical review letters*, 122(19):193904, 2019.

[121] Jinyang Liang, Cheng Ma, Liren Zhu, Yujia Chen, Liang Gao, and Lihong V Wang. Single-shot real-time video recording of a photonic mach cone induced by a scattered light pulse. *Science advances*, 3(1):e1601814, 2017.

[122] Alexander M Puckett, Mark M Schira, Zoey J Isherwood, Jonathan D Victor, James A Roberts, and Michael Breakspear. Manipulating the structure of natural scenes using wavelets to study the functional architecture of perceptual hierarchies in the brain. *NeuroImage*, 221:117173, 2020.

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

[124] David J Field. Relations between the statistics of natural images and the response properties of cortical cells. *Josa a*, 4(12):2379–2394, 1987.

[125] Bruno A Olshausen and David J Field. Vision and the coding of natural images: The human brain may hold the secrets to the best image-compression algorithms. *American Scientist*, 88(3):238–245, 2000.

[126] Geoffrey J Burton and Ian R Moorhead. Color and spatial structure in natural scenes. *Applied optics*, 26(1):157–170, 1987.

[127] Bahaa EA Saleh and Malvin Carl Teich. *Fundamentals of photonics*. john Wiley & sons, 2019.

[128] 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.

[129] 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(4):04LT04, 2021.

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

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

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

[133] 3Blue1Brown. There's more to those colliding blocks that compute pi, March 2025. Accessed: 25-Feb-2025.

[134] Folke Björk, Ola Dickman, and Bengt Stenberg. Long-term studies of rubber materials by dynamic mechanical stress relaxation. *Rubber chemistry and technology*, 62(3):387–425, 1989.

[135] Hao Hu, Zhidong Jia, and Xilin Wang. Aging mechanism of silicone rubber under thermal–tensile coupling effect. *IEEE Transactions on Dielectrics and Electrical Insulation*, 29(1):185–192, 2022.

[136] Kazuyoshi Itoh. Analysis of the phase unwrapping algorithm. *Applied optics*, 21(14):2470–2470, 1982.

[137] Mark Jenkinson. Fast, automated, n-dimensional phase-unwrapping algorithm. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 49(1):193–197, 2003.

[138] Miguel Arevallilo Herráez, David R Burton, Michael J Lalor, and Munther A Gdeisat. Fast two-dimensional phase-unwrapping algorithm based on sorting by reliability following a noncontinuous path. *Applied optics*, 41(35):7437–7444, 2002.

[139] Edouard Berrocal, Elias Kristensson, Mattias Richter, Mark Linne, and Marcus Aldén. Application of structured illumination for multiple scattering suppression in planar laser imaging of dense sprays. *Optics express*, 16(22):17870–17881, 2008.

[140] Elias Kristensson, L Araneo, Edouard Berrocal, J Manin, Mattias Richter, Marcus Aldén, and Mark Linne. Analysis of multiple scattering suppression using structured laser illumination planar imaging in scattering and fluorescing media. *Optics express*, 19(14):13647–13663, 2011.

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

[142] Elias Kristensson, Andreas Ehn, Joakim Bood, and Marcus Aldén. Advancements in rayleigh scattering thermometry by means of structured illumination. *Proceedings of the Combustion Institute*, 35(3):3689–3696, 2015.

[143] Hong-Yi Huang, Zhen-Jia Cheng, Yang Yang, Qing-Yang Yue, and Cheng-Shan Guo. Single-shot ultrafast sequential holographic imaging with high temporal resolution and a large field of view. *Optics Letters*, 44(19):4885–4888, 2019.

[144] Dalong Qi, Fengyan Cao, Shuwu Xu, Yunhua Yao, Yilin He, Jiali Yao, Pengpeng Ding, Chengzhi Jin, Lianzhong Deng, Tianqing Jia, et al. 100-trillion-frame-per-second single-shot compressed ultrafast photography via molecular alignment. *Physical Review Applied*, 15(2):024051, 2021.

Summary of scientific publications

Co-authors are abbreviated as follows:
Vassily Kornienko (VK), Karolina Dorozynska (KD), Elias Kristensson (EK), Yupan Bao (YB)

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

Summary: The main objective of this publication was to investigate the limits of optical multiplexing with FRAME, specifically how the number of multiplexed images affects image quality. This was examined by imaging a rotating computer fan, a high-pressure water injection event, and a stationary resolution target.

Contribution: I designed and built the experimental setup, with guidance from EK. VK assisted with the software integration of the digital mirror device (DMD). I developed the control software for both the DMD and the camera. I conducted all experiments and performed the majority of the data analysis. VK and EK carried out the analysis related to the resolution target and wrote the corresponding sections of the manuscript. I wrote the remaining parts of the manuscript. I did not participate in the review process due to parental leave. All authors reviewed the manuscript.

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

Summary: The objective of this study was to investigate whether FRAME multiplexing of both multispectral illumination and multispectral emission could enhance fluorophore classification. This was examined using doubly modulated FRAME, generating a total of 24 spectral snapshot images of nine different fluorophores. These images were analyzed and used for classification.

Contribution: I developed and performed the analysis for fluorophore classification. All authors reviewed the manuscript.

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

Summary: The objective of this study was to investigate whether FRAME can be combined with background oriented schlieren (BOS) imaging, for higher sensitivity and contrast when imaging transparent targets. This was achieved by studying a flame, using both a LED based setup and a laser based setup.

Contribution: Building on a previous setup, I finalized the LED configuration with assistance from VK and conducted the LED experiments. I developed the schlieren analysis method for FRAME and performed all related analyses. Additionally, I developed the model for estimating the sensitivity of the method. I wrote the manuscript, and all authors reviewed it.

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

Summary: The purpose of this study was to integrate FRAME with a high-speed camera to achieve microscopic, simultaneous multi-temporal-scale imaging. This approach was applied to a high-pressure injection event, producing a long-duration 21 kfps video composed of FRAME triplets at up to 10 Mfps. This enabled the extraction of 2D acceleration vectors for all distinct image features.

Contribution: I performed the uncertainty calculations, establishing the link between acceleration dispersion and average segmentation error. All authors reviewed the manuscript.

Paper V: Ultrafast single-shot imaging of laser induced plasma events in air

Summary: The primary objective of this study was to demonstrate the capability of FRAME for interferometric imaging of transparent events in a microscopic configuration at ultrafast frame rates. To achieve this, we imaged laser-induced plasma events, including filamentation, in air. The results included highly detailed image sequences capturing chaotic breakdown and filament formation.

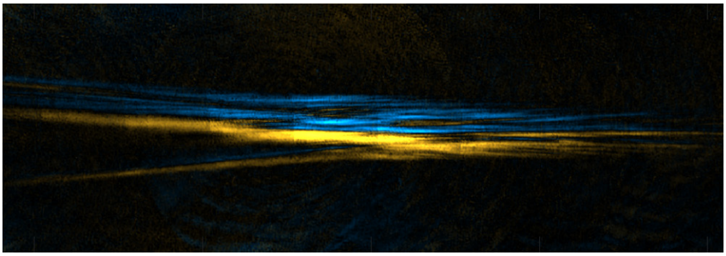
Contribution: I designed, prototyped, and built the 3D-printed unit for probe pulse generation, delay, and modulation. I assembled the experimental setup and developed the control software. I performed the experiments with assistance from YB. I conducted all data analysis and wrote the manuscript, which EK finalized. All authors reviewed the manuscript.

Paper VI: Coherence lifetime imaging

Summary: The objective of this study was to track the molecular dynamics of a laser-induced plasma. This was achieved by combining FRAME with Coherent anti-Stokes Raman Spectroscopy (CARS) to generate an image sequence of laser-induced molecular alignment. The sequence was analyzed to produce a pixel-wise map of the coherence lifetime, linking lifetime to the dynamic state of the process.

Contribution: I designed and assisted in constructing the unit for probe pulse generation, delay, and modulation. I conducted the interferometric analysis of the images.

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



Faculty of Engineering
Department of Physics

Lund Reports on Combustion Physics, LRCP-267
ISBN 978-91-8104-481-2
ISSN 1102-8718