UMR IRISA

Activity Report 2025

Team COMPACT

COMPression of mAssively produCed visual daTa

Joint team with Centre Inria de l'Université de Rennes

D6 – Signal, Image, Language

Université de Rennes | rennes | INSA INSTITUT NATIONAL DES SCIENCES APPLIQUÉES RENNES | cnrs | Inria | Inserm | CentraleSupélec | IMT Atlantique Bretagne-Pays de la Loire École Mines-Télécom | Université Bretagne Sud ubs

Project-Team COMPACT

Creation of the Project-Team: 2024 July 01

Keywords

Computer sciences and digital sciences

A5.9. – Signal processing

A5.9.1. – Sampling, acquisition

A5.9.2. – Estimation, modeling

A5.9.3. – Reconstruction, enhancement

A5.9.4. – Signal processing over graphs

A5.9.5. – Sparsity-aware processing

A5.9.6. – Optimization tools

A8.6. – Information theory

A8.7. – Graph theory

A9.2. – Machine learning

Other research topics and application domains

B3.1. – Sustainable development

B6.5. – Information systems

Contents

Project-Team COMPACT **1**

1 Team members, visitors, external collaborators **4**

2 Overall objectives **5**

3 Research program **7**

4 Application domains **13**

5 Social and environmental responsibility **13**

6 Highlights of the year **14**

7 Latest software developments, platforms, open data **14**
7.1 Latest software developments . . . 14
7.1.1 color-guidance . . . 14
7.1.2 Graph coarsening with message-passing guarantees . . . 14
7.1.3 Taxonomy of reduction matrices for Graph Coarsening . . . 15
7.1.4 mendevi . . . 15
7.2 Open data . . . 15

8 New results **15**
8.1 Axis 1: Compression for specific types of visual data, receivers and media . . . 15
8.1.1 DUALF-D: Disentangled Dual-Hyperprior Approach for Light Field Image Compression . . . 15
8.1.2 Zero-error information theory and application to coding for Computing . . . 16
8.1.3 Coding for Machine: learning in the compressed domain . . . 16
8.1.4 Efficient Constraining of Transcoding in DNA-Based Image Storage . . . 16
8.1.5 Compact image representation for content-based image retrieval in DNA data storage 17
8.1.6 SCALED : Surrogate-gradient for Codec-Aware Learning of Downsampling in ABR Streaming . . . 17
8.1.7 OSLO-IC: On-the-Sphere Learned Omnidirectional Image Compression with Attention Modules and Spatial Context . . . 18
8.2 Axis 2: Sobriety for visual data . . . 18
8.2.1 Semantic compression of images at extremely low bitrate . . . 18
8.2.2 Compressing image encoders via latent distillation . . . 18
8.2.3 Energy-aware images via pixel value reduction: the impact of compression on attenuation maps . . . 19
8.2.4 Experimental analysis of the impact of multi-threading on video encoding energy consumption . . . 19
8.2.5 Efficiency vs sufficiency for video streaming systems . . . 19
8.2.6 Video streaming: how do the socio-economical models shape our research questions? 20
8.3 Axis 3: Acquisition/representation/processing co-design . . . 20
8.3.1 GS-Morph: Dynamic Novel View Synthesis via UDF-ARAP Gaussian Splat Morphing, 20
8.3.2 CAFe-GS: Compactness-Aware Frequency-Guided Densification for 3D Gaussian Splatting . . . 20
8.3.3 Extended-Depth Multispectral Fluorescence Microscopy with Co-Designed Metaoptics and Reconstruction . . . 21
8.4 Axis 4: Learning methods and guarantees . . . 21
8.4.1 MUPET: Maximum A Posteriori Training of Diffusion Models for Image Restoration 21
8.4.2 Taxonomy of reduction matrices for Graph Coarsening . . . 21
8.4.3 Node Regression on Latent Position Random Graphs via Local Averaging . . . 22
8.4.4 Backward Oversmoothing: why is it hard to train deep Graph Neural Networks? . 22

9 Bilateral contracts and grants with industry **23**
9.1 Bilateral contracts with industry . . . 23
9.1.1 CIFRE contract with TyndallFx on Radiance fields representation for dynamic scene reconstruction . . . 23
9.1.2 CIFRE contract with MediaKind on Learned video downscaling for end-to-end Rate-Distortion optimization of video streaming system . . . 23
9.1.3 CIFRE contract with InterDigital on Hybrid conventional and deep learning-based video coding . . . 24
9.1.4 CIFRE contract with InterDigital on End-to-end energy-constrained video content delivery . . . 24

10 Partnerships and cooperations **25**
10.1 European initiatives . . . 25
10.1.1 Horizon Europe . . . 25
10.1.2 H2020 projects . . . 25
10.2 National initiatives . . . 26
10.2.1 PEPR MoleculArXiv. Targeted project 2: From Digital Data to Synthetic DNA . . . 26
10.2.2 PEPR IA. Project SHARP : Sharp Theoretical and Algorithmic Principles for frugal ML . . . 26
10.2.3 ANR Young researcher grant: MAssive multimedia DAta collection REpurposing (MADARE) . . . 27
10.2.4 Joint Project (Défi commun) Nisk.AI . . . 27
10.3 Regional initiatives . . . 28
10.3.1 CominLabs Colearn project: Coding for Learning . . . 28
10.3.2 CominLabs VideoImpact project: Model the environmental cost of video delivery 28
10.3.3 ARED VideoLimit project . . . 29

11 Dissemination **29**
11.1 Promoting scientific activities . . . 29
11.1.1 Scientific events: organisation . . . 29
11.1.2 Journal . . . 30
11.1.3 Invited talks . . . 30
11.1.4 Leadership within the scientific community . . . 30
11.1.5 Scientific expertise . . . 31
11.1.6 Research administration . . . 31
11.2 Teaching - Supervision - Juries - Educational and pedagogical outreach . . . 31
11.2.1 Supervision . . . 31
11.2.2 Juries . . . 32
11.2.3 Internal or external Inria responsibilities . . . 33
11.3 Popularization . . . 33
11.3.1 Specific official responsibilities in science outreach structures . . . 33
11.3.2 Productions (articles, videos, podcasts, serious games, ...) . . . 33
11.3.3 Participation in Live events . . . 33

12 Scientific production **33**
12.1 Major publications . . . 33
12.2 Publications of the year . . . 33
12.3 Cited publications . . . 35

1 Team members, visitors, external collaborators

Research Scientists

- Aline Roumy [Team leader, INRIA, Senior Researcher, HDR]
- Christine Guillemot [INRIA, Senior Researcher, HDR]
- Nicolas Keriven [CNRS, Researcher]
- Natacha Lapeyroux [INRIA, Starting Research Position, from Sep 2025]
- Thomas Maugey [INRIA, Senior Researcher, HDR]

Post-Doctoral Fellows

- Hugo Jaquard [CNRS, Post-Doctoral Fellow, from Mar 2025]
- Caroline Mazini Rodrigues [CNRS, Post-Doctoral Fellow, from Feb 2025]

PhD Students

- Sara Al Sayyed [INRIA]
- Emmanuel Victor Barbosa Sampaio [INTERDIGITAL, CIFRE]
- Stephane Belemkoabga [TYNDAL FX, CIFRE]
- Tom Bordin [INRIA, until Sep 2025]
- Adarsh Jamadandi [CNRS, from Sep 2025]
- Antonin Joly [CNRS]
- Antoine Monier [INTERDIGITAL, CIFRE]
- Esteban Pesnel [MEDIAKIND, CIFRE]
- Remi Piau [INRIA, until Jan 2025]
- Robin Richard [INRIA, from Sep 2025]

Technical Staff

- Robin Richard [INRIA, Engineer, until Aug 2025]

Interns and Apprentices

- Yann Viegas [INRIA, Intern, from Jun 2025]

Administrative Assistant

- Caroline Tanguy [INRIA]

2 Overall objectives

Context

Visual data (images and videos) is omnipresent in various forms (movies, screen content, satellite images, medical images, ...), and provided by different actors ranging from video-on-demand platforms to social networks, and including organizations disseminating Earth observation data. Indeed, video is massively present on the web and accounted for nearly 66% of total internet traffic in 2022 [62]. Therefore, compressing, storing, and transmitting visual data represents a significant societal challenge. Another remarkable fact is that not only does video traffic represent the majority of internet traffic, but it also increases every year. For instance, the number of uploaded hours on Youtube, and shared pictures were mutliplied by 10 and 20 respectively in 10 years (Every minute, 48 hours of videos uploaded in 2013 against 500 hours in 2022, and 3.6K shared pictures in 2013 against 66K in 2022 [40]). This acceleration is predicted to continue. Indeed, video traffic on mobile networks accounted for 71% in 2022 and is predicted to reach 80% by 2028 [42]. To address this issue of ever-increasing data volumes, we analyze the usage of videos more finely, and we realize that within video traffic, we can distinguish between **massively generated data** on one hand and **massively viewed data** on the other hand. Massively generated data can either be provided by machines (for instance, in Copernicus, the Earth observation component of the European Union Space Program, 16 TB of observed or prediction data is provided daily [38]), or humans (in 2022, YouTube saw the upload of 500 hours of video content every single minute [40]). Massively viewed data is mostly movies from video-on-demand platforms. These two modes of traffic have different characteristics, and our team proposes to respond specifically to these two contexts. Finally, another consequence of this massive aspect is the energy and ecological impact associated with the processing, storage, and transmission of this data.

General objective

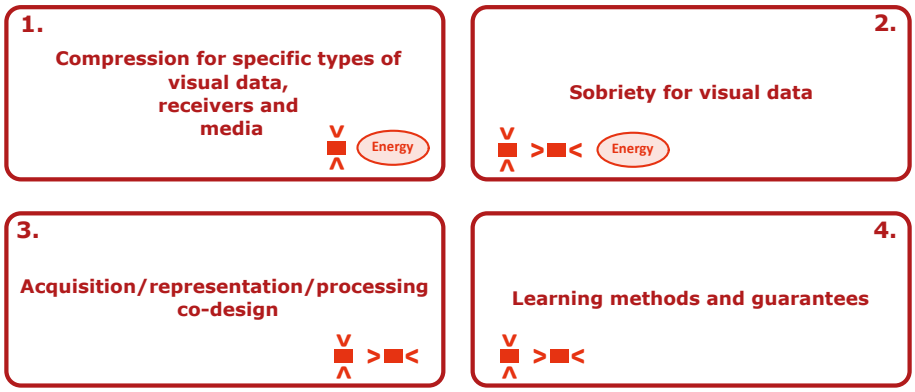
Our main objective is to address the compression problem **in the context of the rapid growth of video usage**, and develop **mathematically grounded algorithms for compressing and processing visual data**. This implies compressing visual data, whose individual volume keeps increasing (new image modalities such as light field, 360, but also higher resolution videos). But it also implies going beyond the classical approach of compressing a single data item to a collection of visual data. To achieve this goal, our team relies on expertise in signal and image processing, statistical machine learning, and information theory. Our originality lies in addressing compression problems in their entirety with contributions that are both practical and theoretical. By doing so, the proposed solutions will address compression challenges comprehensively. More precisely, we begin with a thorough analysis of the compression problem in its practical context taking into account the current context of massively produced data. This will lead to a formulation as an optimization problem and the derivation of information theoretical compression bounds. Subsequently, compression and processing algorithms will be proposed, accompanied by theoretical guarantees regarding content preservation. Finally, validation is performed on real-world data.

Scientific challenges

Compressing this massive data within an ecological transition context leads us to three scientific challenges:

- Reducing the size of **each individual** visual data,
- Reducing the size of a **collection** of visual data,
- Reducing **energy** consumption.

These challenges will be addressed through four main research axes, as shown below:



In the first axis, we will compress data taking into account its usage, i.e., the type of receiver (human versus machine performing inference), as well as its storage mode depending on whether it is hot or cold data. This will both reduce the dimension of the data and provide an energetically efficient solution. In the second axis, the goal is to move towards energy efficiency by proposing algorithms that both reduce the size of individual data and data collections. The third axis also aims to reduce the size of data or data collections, but this time considering the acquisition process and/or a final restoration objective. Finally, many of the proposed methods will be based on machine learning, hence the need to analyze these methods and provide guarantees.

Each axis will be composed of the following sub-axes:

Axis 1. Compression for specific types of visual data, receivers and media,

- **Axis 1.1.** Compression adapted to the data-type,
- **Axis 1.2.** Compression adapted to the user-type: Machine,
- **Axis 1.3.** Compression adapted to the media.

Axis 2. Sobriety for visual data,

- **Axis 2.1.** Ultra-low bitrate visual data compression,
- **Axis 2.2.** Data collection sampling,
- **Axis 2.3.** Low-tech video coders,
- **Axis 2.4.** Sobriety in video usage.

Axis 3. Acquisition/representation/processing co-design,

- **Axis 3.1.** Joint optics/processing,
- **Axis 3.2.** Joint representation/processing: Neural Scene Representation.

Axis 4. Learning methods and guarantees.

- **Axis 4.1.** Optimization methods with learned priors,
- **Axis 4.2.** Learning on graphs,
- **Axis 4.3.** Reducing graphs.

Each of these sub-axes addresses one or several of the initial objectives. Indeed, the first scientific challenge is to **reduce the size of each individual visual data**, such as videos or images. This reduction can be achieved either during acquisition (optics/image processing co-design, compressive acquisition, in Axis 3.1) or after acquisition through a processing leading to a compact representations (low-rank implicit representation, Axis 3.2, learned priors, Axis 4.1, or for a given data type, such as light fields, Axis 1.1). Another approach to size reduction is through the utilization of extremely compact storage mediums, such as DNA storage (Axis 1.3). Furthermore, by considering the usage context, significantly higher compression rates can be achieved when the user is interested in the semantic content rather than the entirety of the

visual data (Axis 2.1), or when performing specific data processing tasks, as in the case of video coding for machines (Axis 1.2).

The second challenge focuses on **reducing the size of a collection of visual data**, for instance, by sampling a database. This sampling can be performed by processing individual data items (Axis 2.2) or by using a structured representation of the database in the form of a graph, addressing issues such as graph reduction (graph sampling, graph coarsening in Axis 4.3), and processing data defined on these graphs (Axis 4.2). Reducing the size of a collection of visual data will also be addressed by learning a compact representation of the whole collection (Axis 3.2).

The third challenge, applicable to both previous challenges, involves **reducing energy consumption**. This will be accomplished through DNA storage research, which offers a low-energy cost storage medium, as well as through optimizing solutions with explicit consideration of global energy costs (for instance in the context of streaming) (Axis 1.3). On top of these necessary efforts for improving the efficiency of coding/storage/transmission systems, a global energy consumption will be targeted, involving the study of efficient and acceptable solutions to aim sobriety in video usage (Axis 2.4).

3 Research program

Axis 1: Compression for specific types of visual data, receivers and media

We start from the observation that visual data is massive but in different ways. For instance, data is individually massive because the dimension of each data point increases, and considering the nature of this data is important for efficient compression (Axis 3). Furthermore, visual data is massively present on networks for different reasons. On one hand, there are massively generated data points that, in some cases, are rarely viewed. On the other hand, there are massively viewed data points that represent a smaller volume than the former. Therefore, it is necessary to propose solutions adapted to each use case.

In the case of massively generated data, the volume of this data is such that it cannot all be visualized by humans. Instead, it will be analyzed by machines, which represents new challenges (Axis 3). Additionally, once analyzed by machines, the rarely viewed cold data can be stored on a medium that allows for low-energy-cost storage, such as DNA (Axis 3). As for the massively viewed data, such as in streaming, the challenge is to offer compression algorithms that optimize not for a financial cost but rather for an energy cost (Axis 3).

Axis 1.1: Compression adapted to the data-type

The field of visual data compression knows new challenges triggered by the emergence of novel modalities (light fields, aka plenoptic , 360^o videos, and even holographic data). This research axis focuses on compact representation of light fields. Unlike traditional cameras which capture simple 2D images, light field cameras capture very large volumes of high-dimensional data containing information about the light rays as they interact with the physical objects in the scene. A major challenge in the practical use of light field technology is the huge amount of captured data, hence the need for efficient compression solutions. While in the past decade the problem has been addressed using traditional signal processing models, e.g. sparse or low rank models, these models present some limitations in terms of well capturing and representing the characteristics of real data. Real data in general require much more complex models that cannot be fully expressed analytically. By contrast, machine learning (ML) methods are data-driven approaches which, by learning a very large number of parameters, turn out to be more powerful for encoding and expressing complex data properties. This is especially important for plenoptic data which represents the complexity of the visual worlds in terms of reflective, diffusive, semi-transparent and partially-occluded objects at various depths. In this context, this research axis aims at dealing with high dimensional light field data, focusing on problems of dimensionality reduction for compression while enabling rendering of high quality. Another problem that will be investigated corresponds to the case where the light field or plenoptic data is first represented by a deep network model. The problem of data compression then becomes a problem of dimensionality reduction of Deep Network Models, e.g. for Mobile Computational Plenoptics.

Axis 1.2: Compression adapted to the user-type: Machine

The volumes of visual data being generated [40] are such that these data will not only be viewed by humans but also by machines. For instance, in autonomous vehicles, the machine is the perception system that processes videos to detect objects such as pedestrians, vehicles, traffic signs, and barriers. Another example is the case when a tremendous amount of visual data is uploaded (in social media for instance) and analyzed to make recommendations to humans. A notable difference between compression for humans and compression for machines is that in the case of machines the entirety of the image is not necessary but only some elements are needed to perform the analysis. Hence there is a need to develop specific algorithms for compression for machines.

Furthermore, among the use cases of compression for machines, we can distinguish two scenarios. In the case of cameras embedded in autonomous vehicles, it is known, upon acquisition, that these visual data will be destined for machines. However, due to time and/or computational constraints, the analysis cannot be performed at the camera, and the data need to be compressed and sent to a remote machine. Instead, in the second example of data uploaded on a social media, the primary destination of the data was initially a human, but it is later decided, after compression, that these data will be analyzed by a machine. For these two use cases, the challenges are different. In the first case, the challenge is to (i) develop new compression algorithms that take into account the receiver, machine and the task that will be performed. In the second case, the goal is instead to (ii) develop algorithms that process the data directly in the compressed domain when the compression algorithm has been specifically designed for human vision.

To develop new compression algorithms (i), our approach is to first define the achievable compression rates when the receiver is a machine that is not interested in the entirety of the data but aims to perform processing on it. Our approach will differ from the work of the community [41, 44, 45] in that we incorporate a strict guarantee on the quality of the processing output. The long term objective is to design compression algorithms, where the task may not be known in advance or another task may be chosen (for instance, a new category to be detected).

When the objective is to build algorithms that allow for processing compressed data with an existing algorithm primarily designed for humans (ii), our approach is to avoid decompressing the data. By avoiding data decompression, it is possible to work with more compact representations of the data. The community avoids this decompression when compression is learned for a specific task (i), as in [64, 36, 37]. Conversely, our objective is to construct these algorithms when the compression is performed by an existing algorithm intended for human viewers.

Axis 1.3: Compression adapted to the media

Storing on DNA Data volume growth has led to a projected data storage requirement of 175 ZB by 2025 [61]. However, the actual data storage capacity currently falls short of this forecast. Furthermore, a significant portion of this data is rarely accessed and is categorized as "cold" data. One potential solution to address these challenges is DNA storage as it offers several advantages, including high data density, extended retention, and low energy cost [35]. Indeed, in terms of data density, DNA can store about 10^{19} bytes per cm^3, enabling the storage of all data generated throughout human history within a 30 cm-sided cube [68]. Regarding retention, DNA can endure for centuries, in contrast to contemporary storage mediums that typically last for decades [68]. Additionally, DNA storage is energy-efficient, since it can be stored at reasonable temperatures, if it is kept away from light and humidity.

Nonetheless, making DNA an efficient storage solution involves overcoming numerous challenges. These challenges encompass:
(i) Data Transformation: convert data into a quaternary code (ACGT). (ii) DNA Synthesis: write data, essentially synthesizing DNA. (iii) DNA Sequencing: extract the quaternary code from DNA, i.e., sequencing DNA. (iv) Data Retrieval: transform back the read quaternary code into the original data. Our primary objective is to address the first and fourth challenges by developing compression algorithms that are robust to synthesis and, more significantly, sequencing errors that occur during steps (ii) and (iii). Indeed, efficient DNA storage heavily relies on rapid sequencing methods, which introduce errors. For instance, real time analysis has been achieved at the price of increased error rates with nanopore sequencing, developed by Oxford Nanopore Technologies (ONT). The main difficulty comes from the type of errors: nanopore introduces not only conventional substitution errors but also unconventional deletion and insertion errors. Deletion

differs from erasure errors, where it is known which part is missing (e.g., lost packets on the internet can be identified by packet headers). Such knowledge of the existence and position of the missing part is unavailable for deletions, and this complicates the correction of this type of error. While the research community largely concentrates on constructing error-correcting codes, our approach aims to develop compression algorithms that are resilient to these errors.

Storing and processing on server for streaming In the case of massively viewed visual data, such as in the case of video streaming, a major objective is to significantly reduce the energy consumption of these solutions. Serving requests is energy-intensive due to the various processing steps undergone by the video before transmission. In fact, the same video content is transmitted with variable qualities (in terms of spatial and temporal resolution, as well as compression errors) in order to adapt to the network bandwidth and receiver type (screen size). In practice, for each request, the high-quality stored video is degraded (in resolution and error level) and then re-compressed. At the decoder level, the video is decompressed and potentially super-resolved to reach the screen resolution. Classically, the optimization of the processing chain is performed to reduce latency and the amount of transmitted data. Instead, our focus is to consider energy consumption as a criterion, and to perform a global optimization taking into account not only transmission, but also storage cost and computation to be performed upon request. This work will be carried out in collaboration with streaming specialist companies. The challenge is to build intermediate representations of videos that provide a video stream compatible with the standard and suitable for transmission (network and screen), thereby optimizing the overall energy balance (storage, server processing, transmission, post-processing at the receiver).

Axis 2: Sobriety for visual data

The sixth report of the Intergovernmental Panel on Climate Change (IPCC) [69] states that if we want to keep the global warming under 1.5°C (Paris agreement), one should target, for 2030, a global emission decrease of 50% when compared to those of 2019. This corresponds to a decrease of 7.6% per year [53]. They also state that *this is not the path that is currently taken*. Hence, every part of our society must urgently aim at sobriety. This is in particular the case of the energy consumed by video data creation/streaming/consumption. In this axis, we will explore solutions enabling a significant reduction of the GreenHouse Gas (GHG) emissions due to video usage. Our strategy is to work on two complementary questions: how to significantly decrease the data size (drastic compression in Axis 3 and data collection sampling in Axis 3)? And how to limit the global video creation and usage (Axis 3)?

Axis 2.1: Ultra-low bitrate visual data compression

The goal of this axis is to reduce the storage cost of cold data, by achieving very high compression ratio. Recently, researchers have proven the existence of a trade-off between distortion and perception when compressing data at low bitrate[32]. In other words, targeting low bitrate inevitably leads to move away from the traditional compression's objective, *i.e.,* keeping faithful decoded data, and to target visual plausibility instead. Therefore, the envisaged solution will semantically describe the visual information in a concise representation, thus leading to drastic compression ratios *exactly as a music score is able to describe, for example, a concert in a compact and reusable form*. This enables the compression to withdraw tremendous amount of useless, or at least not essential, information while condensing the important information into a compact semantic description. At the decoder side, a generative process, relying for example on Diffusion Models [60], is in charge of reconstructing the image or video that is close semantically to the input. In a nutshell, the decoded signals target subjective exhaustiveness of the information description, rather than fidelity to the input data, as in the traditional compression algorithms. Naturally, not all the visual content is meant to be regenerated. Users might be willing to retrieve faithfully the content after decompression. Such approaches will therefore be designed according to user's profile taking into account their choice and interaction. This is a complete change of paradigm, which must enable gigantic compression gains. Considering this approach would use heavy deep learning algorithms and may not tackle data that are often decoded, otherwise the energy due to storage cost reduction would be totally negligible when compared with the huge decoding complexity. On the contrary, this would perfectly fit with cold data. Finally, in order to

be coherent with the purpose of sobriety, we will look for solutions that do not require retraining or even fine-tuning of the heavy Diffusion Models.

Axis 2.2: Data collection sampling

As previously stated, the amount of data created every day is huge *and* exploding. This is certainly accelerated by the fact that most of the social network, video platforms or mobile companies offer the possibility to create, stream and store unlimited data size (or with unreachable bounds), leaving the impression that the storage of data is intangible and cost-less in terms of energy consumption. Increasing the awareness of users or companies requires an efficient way to automatically decide what data deserves to be kept or deleted.

In this axis, we will explore data collection sampling, which consists in selecting the images and videos a user would like to keep among a massive data collection, enabling significant data size savings. This requires first modeling the information perceived by a given user when experiencing a data collection (the initial or the sampled one). This model relies on the volume spanned by the sources features in a personalized latent space. In parallel, we will develop methods to learn the structure and statistics that rule a given data collection. Concretely, among all the pictures of an image collection, some coherent patterns (*e.g.*, landscape, portrait), resemblance between images, chronological landscape evolution or any salient content can be learned and described by mathematical tools, for example with graphs or manifolds. Thereafter these structures will be the support of sampling algorithms aiming at the subjective exhaustiveness of the description, *i.e.*, covering the maximum volume of the learned structure. We will thus pose the trade-off between the rate of the samples (not necessarily taken from the input data, but could be a combination of them) and the quality of the obtained description, driven by the user's preferences.

Axis 2.3: Low-tech video coders

All the recent advances in video compression are due to an increase of the complexity: *e.g.,* more tools and more freedom in the choice of parameters [34] or fully deep learning-based algorithms [55]. In such a context, the global energy cost due to video consumption can only explode, which is not compatible with the urgent need of energetic sobriety. Developing low-energetic video compression/decompression algorithms has been explored for a long time [51, 29, 59]. However, most of the time, the achieved low complexity of the compression algorithms comes from the reduction of the capability of the video coder (*e.g.,* less parameters to estimate, removing of some complex functionalities). Such approaches do not put in question the trade-off between complexity and video coding performance, and thus remain limited.

In this axis, we plan to investigate low complexity algorithms that *are not* low-cost versions of a complex algorithm. The proposed methodology is the following. We start from a complex learning-based coder as for example the auto-encoder-like architecture proposed in [54]. Such architectures are able to achieve outstanding performance, with, however a gigantic encoding and decoding complexity. Our goal is to investigate how to *deduce* from this trained network and its millions of parameters, some efficient features for low complexity compression. As an example, we can show that the set of non-linear operations involved in a deep convolutional neural architecture can be modeled as a linear operation *once the input is fixed*, like it is studied in [58, 57]. The strength of the deep architecture resides in its ability to adjust this linear filter to the input. For our purpose, we will, on the contrary, investigate if some common features reside in these linear filters when the input is changed. These common features may constitute, for example, an efficient transform or partitioning operation that does not require anymore millions of parameters. In a nutshell, the intuition will be to take benefit of algorithms trained on a large set of images and to extract from them some common analysis tools.

Axis 2.4: Sobriety in video usage

Rebound effect or the Jevons's paradox [67] refers to the fact that reducing the cost (in terms of energy or resource consumption) of a technology often leads to an increase of the technology usage and thus to a global increase of the cost, in opposition with the initial goal. Video compression is clearly a good example of this rebound effect. Smaller video sizes (and other technology advances) have led to a global increase of the video usage in today's society. As the ultimate goal, for achieving IPCC objectives, is to reduce the global carbon footprint of video usage, compression nowadays should not only focus on the reduction of each

video file individually. The compression problem should be formulated globally. This inevitably raises the following research question: what is the best (most efficient and acceptable) solution for reducing the amount of videos created/stored/consumed? This question naturally includes the study of user's behavior, and thus deals with other research fields in human and social sciences. The goal of the team COMPACT is twofolds: i) to raise a multidisciplinary research effort on that question by connecting different laboratories and ii) to put its expertise in video compression to the service of this crucial question.

Axis 3: Acquisition/representation/processing co-design

In this axis, the goal is to compress either a data or a collection of data, while taking into account either the acquisition process or a final restoration objective.

Axis 3.1: Joint optics/processing

Our goal is the design of an end-to-end optimization framework designed for acquiring high-resolution images across an extensive Depth of Field (DOF) range within a microscopy system. Microscopy is indeed one key potential application of light field imaging. The optics and post-processing algorithm will be modeled as parts of the end-to-end differentiable computational image acquisition system, allowing for simultaneously optimizing both components. Our computational Extended DOF microscopy imaging system will employ a hybrid approach combining an optical setup with a learned wavefront modulating optical element at the Fourier plane based on metasurfaces. The extended depth of field leads to an increased axial resolution which refers to the ability to distinguish features at different depths by refocusing. While we have obtained initial results for 2D microscopy [30], our goal here will be to extend these results to light field microscopy, which has recently retained the attention of the research community [63, 65, 52].

Axis 3.2: Joint representation/processing: Neural Scene Representation

The task of generating high-quality immersive content with a sufficiently high angular and spatial resolution is technologically challenging, due to the complexity of the constrained capture setup and the bottleneck of data storage and of computational cost. Reconstructing the imaged scene (from a few viewpoints), with a sufficient resolution and quality, and in a way that we can observe it from almost continuously varying positions or angles in space is also an important challenge for a wide adoption in consumer applications. To address the two above problems, the concept of NeRF has been introduced as an implicit model that maps 5D vectors (3D coordinates plus 2D viewing directions) to opacity and color values. The model is based on multi-layer perceptrons (MLP) trained by fitting the model to a set of input views. The learned model is an implicit scene representation that can be used to generate any view of the light field using volume rendering techniques. A variety of works have attempted to handle dynamic scenes in radiance field reconstructions but they either constrain the capture process with multi-view or suffer from quality loss when compared to static scene representations. The proposed research, jointly addressing acquisition, representation and scene reconstruction problems [47, 56, 43] will focus on the reconstruction of neural radiance fields from a limited set of input images, especially in the context of unconstrained, monocular captures, on the completion of the NeRF representation when the capture is incomplete due to a limited set of input images or due to motion in the scene, on the representations of dynamic scenes that are both compact (low memory) and limited in computational complexity. The compactness of scenes will be explored considering joint implicit representations for a collection of data points (2D Images or light fields). The implicit representations inspired from the NeRF concept can be seen as neural network based data representations. The generalization of joint implicit representations to unseen data points assumed to reside in the same subspace as the training data points will also be investigated.

Axis 4: Learning methods and guarantees

A difficulty in visual data (image and video) processing is that their distribution is not known. Therefore, learning-based methods have a certain advantage over model-based methods because they can better adapt to this data. We propose to explore two new ideas in the context of these learning-based methods with the goal of obtaining guarantees on the quality of processing. First, in the context of inverse problems, where the

dimension of the observed data is lower than that of the data to be restored, we wish to study the construction of learned priors rather than handcrafted ones, with guarantees stemming from a technique called Deep Equilibrium. In a second approach, we aim to exploit the data's structure (such as a graph), build new learning algorithms adapted to this structure, and obtain theoretical guarantees regarding the learning of the graph but also the learning on the constructed graph.

Axis 4.1: Optimization methods with learned priors

Building upon our past work aiming at taking advantage of learned priors in optimization algorithms, i.e. via plug-and-play and unrolled optimization methods, we will further investigate Deep equilibrium (DEQ) [31] models. Unrolled optimization methods, by coupling optimization algorithms with end-to-end trained regularization, recently emerged as powerful solutions to inverse problems. However, training such unrolled neural networks end-to-end can come with a large memory footprint [46], hence their numbers of iterations are in general limited and they do not generally converge. DEQ models can be seen as an extension of unrolled methods with a theoretically infinite amount of iterations. DEQ models leverage fixed-point properties, allowing for simpler back-propagation. We will further study these models to learn image priors and apply them to inverse problems in classical 2D and new imaging modalities (light fields, omni-directional images).

Axis 4.2: Learning on graphs

In the last decades, there has been a multiplication of data that cannot be properly represented by conventional means, but rather by *relationships* between objects, of various natures and with various properties. Such structures are usually represented as *graphs*. This is for instance the case of collections of (visual) data under the form of *relational databases* (Axis 3), formed by drawing "meaningful" relations between individual data points according to some notion of proximity (semantic, geographical, etc.). Moreover, graphs are increasingly used to represent the structure of (potentially pre-trained) neural networks. Processing this structure using graph machine learning and graph signal processing tools gives rise to the recent topic of *(graph) meta-networks* [49], which draws connections with all other axes, particularly the definition of low-tech encoders (Axis 3). Finally, graphs are also a popular representation for geometric data exhibiting *invariance* to certain transforms [33] such as 2D or 3D isometries, often encountered in non-conventional visual data (Axis 3).

(Un)structured data such as graphs posit many challenges. Processing and storing them can be computationally burdensome if done naively. The main challenge resides in the fact that the *regularity* of other types of data (fixed-size vectors, regular grids, well-defined boundaries, etc.), at the basis of many methods, cannot be easily defined here. This axis is thus dedicated to advancing the state-of-the-art in processing efficiently graph data, often through the lens of compression. ML techniques have proved extremely efficient in designing adaptive, data-driven methods *for* compression [66], including for database reduction [39]. Conversely, the extraction of information *from* compressed databases, *a fortiori* by ML, is a major requirement of any compression pipeline. Since graphs have become the *de facto* structure to represent modern relational data, *graph ML* (GML) has known a tremendous development in the last few years, with Graph Neural Networks (GNN) at the forefront of it. Acclaimed for their flexibility, these deep architectures however suffer from many issues, with very limited theoretical and empirical comprehension. A major goal will be to deepen this understanding through the use of tools such as statistical models of large random graphs and information theory. New random graph models adapted to modern real-world data will be developed, focusing on databases arising from visual data but also generic databases, whose analysis will help the choice of GNN architecture, and ultimately lead to new architecture improving the state-of-the-art, in terms of performance and/or computational efficiency.

Axis 4.3: Reducing graphs

Data compression approaches on graphs are referred to as *graph reduction* methods. With modern large graphs numbering millions of nodes, these methods have become a staple of many pipelines, including ML methods mentioned above and database reduction (Axis 3). Graph reduction can be broadly sorted into two related families of algorithms: graph sampling, and graph coarsening.

Graph sampling Graph sampling consists in selecting, often randomly, a reduced number of "representative" node from a large graph. The means to do so, and the downstream tasks to achieve with the subsampled graph, can take many different forms. Particularly interesting for us is the role of graph sampling for fast and efficient querying in large databases [48] (Axis 3), and reducing the size of large neural networks (Axis 3). We will focus on theoretically grounded methods using models of random graphs and information theory, taking into account the specificity of the graph data examined through the previous axes. Since graph sampling is also part of several modern architectures of GNNs, we will incorporate our methods in such models, and examine in which measure sampling methods can be adaptive, data-driven, and/or trained in an end-to-end manner, taking inspiration from modern generative models. Validation will be performed along different criteria, focusing on the classical trade-off between compression rate and performance score, with different choices for the latter depending on the application: supervised classification accuracy, clustering coefficient, etc.

Graph coarsening A related, but somewhat more complex and less well-defined, problem to graph sampling is that of *graph coarsening*, that is, producing an entirely new smaller graph from a large given graph. Again, the purposes can be many, and graph coarsening has an important role in many efficient methods to query and store large databases [50]. Traditional graph coarsening methods seek to preserve certain property of the graph, e.g. spectral properties, and build specific loss functions and performance measurements around these notions.

We will examine whether different coarsening criteria could be defined in a task-dependent manner with guarantees, for instance with the purpose of reducing large neural networks with graph meta-networks [49] (Axis 3), or to expressly design well-adapted convolution operators to be incorporated in neural nets acting on non-Euclidean data (Axis 3). On the theoretical side, we will examine if additional regularity under the form of random graphs models can be exploited. An information-theoretical approach could also lead to new methods. Moreover, graph coarsening is at the heart of *pooling* in GNNs, a very promising lead to improving such architectures by making them "hierarchical" like CNNs, which is still largely open despite an extensive literature on the topic. A more theoretically-grounded approach to the problem could lead to significant advances in this domain.

4 Application domains

Our research is inherently motivated by the application of image and video compression and processing (mostly to help compression denoising, extrapolating such as super-resolution, view synthesis; in the case, of communication to machine, the final goal of object detection and tracking will be also considered, but here as to measure the efficiency of the compression). Two major types of visual data will be considered. First, hot data, such as publicly available data commonly streamed. We will also consider cold data, such as the archival of data that is rarely accessed, as in the case of legal repositories.

5 Social and environmental responsibility

Most of the research fields tackled by the COMPACT team, such as image/video compression, data dimensionality reduction, are inherently aligned with the objective of bringing frugality for processing algorithms. In other words, our algorithms are designed to reduce the energy and resources required for data analysis and consumption. However, while crucial, this research goal is not sufficient to achieve an effective reduction of the environmental footprint of the digital world.

Indeed, the well-known rebound effect makes that such reductions at the algorithm level implies an increase at a broader level (e.g., more videos being created, more learning models being deployed, etc.). The COMPACT team is well aware of this challenge, and is therefore making a strong effort to build collaborations with Social and Human Science researchers. This interdisciplinary approach aims to explore to what extend some limits in the technology usage may be set.

6 Highlights of the year

In 2025, the team achieved several notable results, including publications in flagship conferences and leading journals in the field, as well as distinguished awards. Noteworthy examples include:

- a study on reduction matrices for graph coarsening [13], published at NeurIPS;
- a contribution to zero-error information theory [5], published in the IEEE Transactions on Information Theory;
- and work on view synthesis, for which Stéphane Belemkoabga was runner-up for the Best Paper Award at the CVMP conference for the paper [16] - [post]

In addition, 2025 was marked by a strong collaboration with InterDigital, initiated in the context of a joint research challenge (défi commun Nisk.AI).

Several new research directions were also launched. In particular:

- research on DNA data storage was initiated and led to first publications, along with the delivery of a tutorial in the framework of the MoleculArXiv Autumn School on DNA Data Storage;
- the COMPACT team initiated a pluridisciplinary collaboration with Social and Human Sciences. This effort was made possible through the CominLabs project “VideoImpact”, which supported the recruitment of Natacha Lapeyroux (sociologist) and fostered collaborations with economists (LEGO laboratory, IMT Brest) and sociologists (ARENES laboratory at Univ Rennes and UCO, Nantes).

7 Latest software developments, platforms, open data

7.1 Latest software developments

7.1.1 color-guidance

Keyword: Image compression

Scientific Description: This study addresses the challenge of controlling the global color aspect of images generated by a diffusion model without training or fine-tuning. We rewrite the guidance equations to ensure that the outputs are closer to a known color map, without compromising the quality of the generation. Our method results in new guidance equations. In the context of color guidance, we show that the scaling of the guidance should not decrease but rather increase throughout the diffusion process. In a second contribution, our guidance is applied in a compression framework, where we combine both semantic and general color information of the image to decode at low cost. We show that our method is effective in improving the fidelity and realism of compressed images at extremely low bit rates (0.001 bpp), performing better on these criteria when compared to other classical or more semantically oriented approaches.

Functional Description: Official implementation of the article: "Linearly transformed color guide for low-bitrate diffusion based image compression" Paper(https://arxiv.org/pdf/2404.06865)

Publication: hal-04882103

Contact: Tom Bordin

7.1.2 Graph coarsening with message-passing guarantees

Keywords: Graph Neural Networks, Deep learning, Dimensionality reduction

Functional Description: This repository contains the code for the paper “Graph coarsening with message-passing guarantees”, published at NeurIPS 2024.

This code includes Jupyter notebooks that reproduce the results (tables and plots) presented in the paper. These experiments focus on using a newly proposed Propagation matrix for the Graph Neural Network (GNN) on the coarsened graph.

Publication: hal-04617519

Contact: Antonin Joly

7.1.3 Taxonomy of reduction matrices for Graph Coarsening

Keywords: Deep learning, Dimensionality reduction, Graph Neural Networks

Functional Description: This repository contains the code for the paper "Taxonomy of Reduction Matrices for Graph Coarsening", published at NeurIPS 2025.

This code includes Jupyter notebooks that reproduce the results (tables and plots) presented in the paper. These experiments focus on optimizing reduction matrices for a fixed lifting matrix in graph coarsening with the framework described in the paper.

Publication: hal-05248172

Contact: Antonin Joly

7.1.4 mendevi

Name: Energy measurement of video encoding and decoding

Keywords: Energy, Video analysis, Video compression

Functional Description: 1. It supports the libx264, libopenh264, libx265, libvpx-vp9, libaom-av1, libsvtav1, librav1e and vvc cpu encoders. 2. It supports the h264_nvenc, hevc_nvenc, av1_nvenc and *_vaapi gpu encoders. 3. Distortions are measured using the lpips, psnr, ssim, vif and vmaf metrics. 4. Complexity are measured using the rms_sobel and rms_time_diff metrics. 5. Encoding efforts are fast, medium and slow. 6. It takes care about the colorspaces (range, transfer and primaries). 7. Iterate over different effort, encoder, mode, quality, threads, fps, resolution and pix_fmt. 8. Energy measurements are catched with RAPL and an external wattmeter on grid'5000. 9. Get the cpu, gpu, ram and temperature activity. 10. Get a full environment context, including hardware and software version. 11. It support the mode (constant bitrate) cbr and (constant quality) vbr. 12. Ability to modify ffmpeg commands on the fly to perform specific tests. 13. It take care to transfer files to RAM if possible to avoid biases related to storage space access. 14. Provides a guide to compile ffmpeg with all optimizations in order to compare encoders/decoders at their limits.

URL: https://mendevi.readthedocs.io/latest/

Contact: Robin Richard

7.2 Open data

8 New results

8.1 Axis 1: Compression for specific types of visual data, receivers and media

8.1.1 DUALF-D: Disentangled Dual-Hyperprior Approach for Light Field Image Compression

Participants: Soheib Takhtardeshir, Christine Guillemot.

Light field (LF) imaging captures spatial and angular information, offering a 4D scene representation enabling enhanced visual un- derstanding. However, high dimensionality and redundancy across spatial and angular domains present major challenges for com- pression, particularly where storage, transmission bandwidth, or processing latency are constrained. We have developed a novel Variational Autoencoder (VAE)-based framework that explicitly disentangles spatial and angular features using two parallel latent

branches [17], [9]. Each branch is coupled with an independent hyperprior model, allowing more precise distribution estimation for entropy coding and finer rate-distortion control. This dual-hyperprior structure enables the network to adaptively compress spatial and angular infor- mation based on their unique statistical characteristics, improving coding efficiency. To further enhance latent feature specialization and promote disentanglement, we introduced a mutual information-based regularization term that minimizes redundancy between the two branches while preserving feature diversity. Unlike prior methods relying on covariance-based penalties prone to collapse, our information-theoretic regularizer provides more stable and interpretable latent separation [8]. Experimental results on publicly available LF datasets demonstrate our method achieves strong compression performance, yielding an average BD-PSNR gain of 2.91 dB over HEVC and high compression ratios (e.g., 200:1). Additionally, our design enables fast inference, with a total end-to- end time over 19x faster than the JPEG Pleno standard, making it well-suited for real-time and bandwidth-sensitive applications. By jointly leveraging disentangled representation learning, dual-hyperprior modeling, and information-theoretic regularization, our approach offers a scalable, effective solution for practical light field image compression.

8.1.2 Zero-error information theory and application to coding for Computing

Participants: Aline Roumy.

Zero-error coding encompasses a variety of source and channel coding problems in which the probability of error must be exactly zero. This requirement is stricter than that of the classical vanishing-error regime, where the error probability tends to zero as the code blocklength goes to infinity. An example of a zero-error problem is coding for computing, where the goal is to compress data not merely for visualization, but also to enable reliable inference tasks.

In general, zero-error coding leads to challenging open combinatorial problems. In [5], we investigated two unsolved zero-error settings: the source coding problem with side information and the channel coding problem. We focused on families of independent problems for which the underlying probability distribution decomposes as a product of marginal distributions. A crucial step in our analysis was establishing the additivity of the optimal rate. Unlike in the vanishing-error regime, this property does not always hold in the zero-error setting. When additivity does hold, concatenation of optimal codes remains optimal.

As a consequence, we derived new single-letter characterizations of the optimal information-theoretic rates for previously unsolved graph families. In particular, we obtained results for graphs formed as products of perfect graphs (which are not perfect in general) as well as for graphs obtained as the product of a perfect graph and the pentagon graph.

8.1.3 Coding for Machine: learning in the compressed domain

Participants: Rémi Piau, Thomas Maugey, Aline Roumy.

In most of the learning tasks, it is necessary to scale the image size to the networks. This downsampling is generally done in the pixel domain (it can be done before or inside the network itself) and thus requires a decoding of the image at its full resolution which can be complex for the most recent formats. Instead, we proposed to sample the image directly in the JPEG bitstream, to partially decode some image MCU and to feed them to the learning task, which is challenging due to the variable length coding involved in JPEG. After showing some interesting properties of the JPEG bitstream, we proposed an end-to-end learning pipeline starting from a decoding of only a extracted subset of the JPEG bitstream. Our results demonstrated the validity of our approach and that learning directly in the JPEG bitstream is possible. [25]

8.1.4 Efficient Constraining of Transcoding in DNA-Based Image Storage

Participants: Sara Al Sayyed, Aline Roumy, Thomas Maugey.

DNA has emerged as a promising alternative for long-term data storage due to its high capacity, durability, and low-energy potential. However, storing data in DNA presents several challenges. First, it requires complex and costly biochemical processes, making efficient compression crucial to reducing DNA synthesis time and cost. Second, these processes are prone to errors that must be avoided and/or corrected. In particular, homopolymers (repetitions of the same nucleotide) are a well-known source of errors during the sequencing step. Avoiding such repetitions helps mitigate errors but introduces a constraint that may increase the data compression rate. In this paper, we propose two transcoding methods that address these two key challenges: reducing data rate and minimizing errors. The first method strictly enforces the error-minimization constraint by eliminating homopolymers of a certain length, at the cost of an increased data rate. In contrast, the second method accepts a slight increase in homopolymers. However, we show that these increases remain limited (2.14% increase in compression rate for the first method and 0.39% homopolymer rate for the second). These two approaches demonstrate that it is possible to efficiently constrain transcoding while balancing error minimization and compression performance. This work was published in [10].

8.1.5 Compact image representation for content-based image retrieval in DNA data storage

Participants: Sara Al Sayyed, Aline Roumy, Thomas Maugey.

In this work, we propose a novel image compression method for content-based image retrieval in the context of DNA data storage. As explained before, storing data on DNA is an extremely promising solution due to its compactness, long-term durability, and energy efficiency. However, its compactness introduces two challenges: the need for efficient data access and the ability to flexibly handle new (and not predefined) types of queries. To address the efficiency challenge, our approach enables direct image retrieval within the DNA domain. To ensure flexibility, we design a compact data identifier that is a semantic representation of the image and serves as a header at the beginning of the DNA strand. Our approach shows high visual and quantitative performance, outperforming state-of-the-art method for various types of query. This highlights that hybridization can be effectively modeled using cosine similarity, without the need for training. This work was published in [11].

8.1.6 SCALED : Surrogate-gradient for Codec-Aware Learning of Downsampling in ABR Streaming

Participants: Esteban Pesnel, Aline Roumy, Thomas Maugey.

The rapid growth in video consumption has intro- duced significant challenges to modern streaming architectures. Over-the-Top (OTT) video delivery now predominantly relies on Adaptive Bitrate (ABR) streaming, which dynamically adjusts bitrate and resolution based on client-side constraints such as display capabilities and network bandwidth. This pipeline typically involves downsampling the original high-resolution content, encoding and transmitting it, followed by decoding and upsampling on the client side. Traditionally, these processing stages have been optimized in isolation, leading to suboptimal end-to-end rate-distortion (R-D) performance. The advent of deep learning has spurred interest in jointly optimizing the ABR pipeline using learned resampling methods. However, training such systems end-to-end remains challenging due to the non-differentiable nature of standard video codecs, which obstructs gradient-based optimization. Recent works have addressed this issue using dif- ferentiable proxy models, based either on deep neural networks or hybrid coding schemes with differentiable components such as soft quantization, to approximate the codec behavior. While differentiable proxy codecs have enabled progress in compression-aware learning, they remain approximations that may not fully capture the behavior of standard, non-differentiable codecs. To our knowledge, there is no prior evidence demonstrating the inefficiencies of using standard codecs

during training. In this work, we introduce a novel framework that enables end-to- end training with real, non-differentiable codecs by leveraging data-driven surrogate gradients derived from actual compression errors. It facilitates the alignment between training objectives and deployment performance. Experimental results show a 5.19improvement in BD-BR (PSNR) compared to codec-agnostic training approaches, consistently across the entire rate-distortion convex hull spanning multiple downsampling ratios. This work was published in [15].

8.1.7 OSLO-IC: On-the-Sphere Learned Omnidirectional Image Compression with Attention Modules and Spatial Context

Participants: Thomas Maugey.

Developing effective 360-degree (spherical) image compression techniques is crucial for technologies like virtual reality and automated driving. This work advances the state-of-the-art in on-the-sphere learning (OSLO) for omnidirectional image compression framework by proposing spherical attention modules, residual blocks, and a spatial autoregressive context model. These improvements achieve a 23.1% bit rate reduction in terms of WS-PSNR BD rate. Additionally, we introduce a spherical transposed convolution operator for upsampling, which reduces trainable parameters by a factor of four compared to the pixel shuffling used in the OSLO framework, while main- taining similar compression performance. Therefore, in total, our proposed method offers significant rate savings with a smaller architecture and can be applied to any spherical convolutional application. This work was published in [18].

8.2 Axis 2: Sobriety for visual data

8.2.1 Semantic compression of images at extremely low bitrate

Participants: Tom Bordin, Thomas Maugey.

We propose a framework for semantic image compression targeting ultra-low bitrates (~0.001 bpp). The semantic content of an image is transmitted through its representation in the CLIP embedding space. Although embeddings lack positional information, semantic features provide strong priors that can be modeled with attention layers (instead of color map as introduced in previous work). We leverage these priors to transmit only residual positional data as attention maps, thereby correcting the spatial arrangement of objects in the scene. Our method is evaluated using both standard objective metrics and subjective human assessments, demonstrating state-of-the-art performance in both aspects. This work is currently under review.

However, in applications targeting extremely low bitrates (0.01 bpp), where the reconstruction distortion can be severe, it makes sense to prioritize parts of the image that are more relevant than others. In a second work, we propose a semantic compression framework that integrates user or application preferences to compress image parts based on their semantic representation. We design a guide for trained diffusion models that takes into account the preferences for describing objects with varying accuracies. We show that we are able to preserve the selected objects while also preserving the semantic and global aspect of the image without any retraining or fine-tuning. This work is currently under review.

8.2.2 Compressing image encoders via latent distillation

Participants: Caroline Mazini-Rodrigues, Nicolas Keriven, Thomas Maugey.

Deep learning models for image compression often face practical limitations in hardware-constrained applications. Although these models achieve high-quality reconstructions, they are typically complex, heavyweight, and require substantial training data and computational resources. We propose a methodology

to partially compress these networks by reducing the size of their encoders. Our approach uses a simplified knowledge distillation strategy to approximate the latent space of the original models with less data and shorter training, yielding lightweight encoders from heavy-weight ones. We evaluate the resulting lightweight encoders across two different architectures on the image compression task. Experiments show that our method preserves recon- struction quality and statistical fidelity better than training lightweight encoders with the original loss, making it practical for resource-limited environments. This work is currently under review [28].

8.2.3 Energy-aware images via pixel value reduction: the impact of compression on attenuation maps

Participants: Emmanuel Sampaio, Thomas Maugey.

Video consumption accounts for a significant share of global energy use, with end devices responsible for most of it. On end devices, display technology plays an important role in energy consumption. Interestingly, OLED technology allows power to be adapted via pixel-intensity manipulation. In this context, Pixel Value Reduction (PVR) has shown promising results for lowering display power by generating attenua- tion maps that adapt image luminance. However, the use of this technology in streaming services has not been fully studied. In this work, we analyze the effect of attenuation-map compression on perceptual quality, bitrate overhead, and end-device energy consumption. Using a pixel-value- reduction model, we generate attenuation maps for target power-reduction levels (10%, 20%, and 40%) and encode them with the HEVC video codec at various quantization- parameter (QP) values (i.e., codec QP). Experiments on 4K content with real OLED power measurements show that compressed attenuation maps maintain high fidelity to the originals, achieving different levels of power reduction with negligible quality loss. Moreover, the results indicate that proper alignment between content and map quantization pa- rameters is critical for reducing bitrate overhead. These findings highlight the feasibility of transmitting compressed attenuation maps to minimize display's energy consumption. This work is currently under review.

8.2.4 Experimental analysis of the impact of multi-threading on video encoding energy consumption

Participants: Robin Richard, Thomas MaugeyCarCar.

Modern CPUs are equipped with more and more cores, raising the question of how parallelism leads to better energy efficiency, especially in intensive tasks like video encoding. This work investigates how video encoding using multiple threads leads to better usage of available cores, and if it actually improves energy efficiency. Based on real video transcoding energy measurements on a server, we test classical energy models in a multi-threaded context. On the one hand, we observe that the energy consumed during encoding is indeed decreasing with the number of cores used during the task. On the other hand, we also observe that this number of used cores is not always linked to the number of threads that are given in parameter to the encoder. Hence, this study enables to state that the energy savings due to multi-threading is likely for small number of threads, but less achievable when the number of threads becomes too high. This work is currently under review.

8.2.5 Efficiency vs sufficiency for video streaming systems

Participants: Thomas Maugey, Anne-Cécile Orgerie, Robin Richard.

To reduce the ecological impact of a technology, scientists often focus on energy efficiency issues, ignoring the complex rebound effects generated by efficiency. We focus on the video transmission technology, and discuss the urgent need to be able to set limits in order to target absolute sustainability and sufficiency.

We show that these limits can provoke opposition or circumvention, illustrating the difficulty of the task. We conclude that the question of limits must be considered as a research problem in its own right, and that it is intrinsically multidisciplinary. This work has been presented in [22].

8.2.6 Video streaming: how do the socio-economical models shape our research questions?

Participants: Natacha Lapeyroux, Thomas Maugey, Anne-Cécile Orgerie.

According to a various number of studies, the environmental and social impacts of video streaming is huge and growing. Today, the work of researchers in the field of image processing only accelerates this explosion by contributing to the emergence of new technologies. At best, researchers are simply trying to improve the efficiency of streaming systems, which, due to the rebound effects, also contributes to "accelerating the acceleration". In this talk, we give an overview of the socio-economical models ruling most of the video streaming platforms, and we show that the research questions tackled nowadays are directly shaped from these models. We also show that these models irremediably lead to bigger videos and more videos. Tackling the reduction of video streaming impacts will only be possible by questioning these models

8.3 Axis 3: Acquisition/representation/processing co-design

8.3.1 GS-Morph: Dynamic Novel View Synthesis via UDF-ARAP Gaussian Splat Morphing,

Participants: Stephane Belemkoabga, Christine Guillemot, Thomas Maugey.

Monocular view synthesis in dynamic scenes remains a fundamental challenge in vision and graphics, particularly for applications like augmented reality, virtual production, and free-viewpoint video. Recovering accurate 3D geometry and realistic rendering from a single RGB-D stream is highly ill-posed due to partial, noisy, and temporally inconsistent observations under non-rigid motion. Recent methods, such as dynamic NeRFs and 4D Gaussian Splatting, attempt to jointly optimize motion and geometry. While effective near training trajectories, these entangled designs often struggle to generalize across novel views and times. We introduce a new framework that explicitly decouples geometry reconstruction and motion estimation to improve robustness and generalization. Given a monocular RGB-D sequence with known poses, we first extract per-frame point clouds and estimate frame-to-frame deformation fields using Unsigned Distance Field (UDF) registration with ARAP regularization. These are used to segment the sequence into motion-coherent Groups of Pictures (GoPs). Each GoP undergoes alternating fusion and deformation propagation to yield a consistent local geometry and dense deformation field. GoPs are then hierarchically merged into a global scene model with a unified deformation field. A spatio-temporal 3D Gaussian Splatting representation is initialized from this model and further refined with photometric and geometric losses. To evaluate generalization, we introduce a two-level protocol: Level 1 tests novel views along the training path, while Level 2 tests novel views at unseen times or poses. We also release a new RGB-D dataset for monocular dynamic scene recon- struction. Our method sets a new state-of-the-art, outperforming prior work in both synthesis quality and deformation accuracy. This work was published in [16].

8.3.2 CAFe-GS: Compactness-Aware Frequency-Guided Densification for 3D Gaussian Splatting

Participants: Christine Guillemot, Leo-Paul Huar.

3D Gaussian Splatting (3DGS) represents scenes using Gaus- sian primitives and enables real-time novel view synthesis. Adaptive Den- sity Control (ADC), a key part of the pipeline, governs when to den- sify these primitives to balance reconstruction quality and efficiency. In the original 3DGS pipeline, densification is triggered by a thresholded positional-gradient criterion. However, this criterion frequently selects already

well-covered regions, leading to redundant primitives and pro- viding weak control over the balance between reconstruction quality and compactness (i.e., fidelity versus primitive count). In CAFe-GS, we pro- pose a new densification criterion based on a per-Gaussian score obtained by mapping per-pixel rendering errors back to the contributing primi- tives, using their effective-opacity under front-to-back alpha composit- ing as weights. The score is then modulated by frequency guidance derived from Laplacian-of-Gaussian responses, promoting detail-rich, high- frequency areas in contrast to smooth or already well-reconstructed re- gions. This criterion drives densification through standard cloning and splitting operations. CAFe-GS provides a clearer, single-parameter han- dle on the quality–compactness balance. Experiments on standard benchmarks show that CAFe-GS achieves comparable PSNR using 2 to 4 times fewer Gaussians at matched quality, and up to 12 to 15 times fewer Gaussians at a controlled PSNR trade-off.

8.3.3 Extended-Depth Multispectral Fluorescence Microscopy with Co-Designed Meta-optics and Reconstruction

Participants: Ipek Anil Atalay Appak, Christine Guillemot.

Fluorescence microscopy can deliver high-resolution spatial details; however, it suffers from shallow depth of field and chromatic aberrations. The impact is greatest for thick specimens and for multispectral data that must stay aligned across depth. We have designed MANTIS (Multispectral All-Depth meta-opTics Imaging System), a co-designed optical–computational platform that achieves extended depth of field from a single acquisition per field of view without axial scanning. A learned meta-optic and a physics-guided reconstruction are trained end-to-end so that depth and wavelength-dependent blur is encoded in a recoverable form and decoded. We target extended depth ranges reaching up to 75 micrometer. The reconstructions show weak depth dependence and low cross-spectral variance. In simulation at 50 micrometer depth of field, mean peak signal-to-noise ratio and structural similarity reach 23.5 dB and 0.70, averaged over depths and channels. We have validated experimentally the designed system by fabricating the learned meta-optic, measuring the point spread functions across the target depths and wavelengths, and reconstructing three-dimensional fluorescence samples. The experimental reconstructions maintain contrast and lateral sharpness across depth, exhibiting modest per-channel variation in PSNR and SSIM, with trends that match the simulation and are consistent with low chromatic aberration and extended depth of field.

8.4 Axis 4: Learning methods and guarantees

8.4.1 MUPET: Maximum A Posteriori Training of Diffusion Models for Image Restoration

Participants: Christine Guillemot, Samuel Willingham.

Inverse problems involve reconstructing clean images from degraded observations. Maximum a Posteriori (MAP) estimation reconstructs the most probable source image from noisy measurements. When combined with Plug-and-Play (PnP) priors defined by an image denoising algorithm, MAP estimation yields high-quality reconstructions. In contrast, Diffusion Models (DMs) address inverse problems by sampling from the posterior distribution using score functions trained on images perturbed by Gaussian noise. Prior work reformulated diffusion sampling as Deep Equilibrium (DEQ) models but did not fine-tune DMs for inverse problems. We have proposed MaximUm a PostEriori Training (MUPET), a framework that leverages PnP gradient descent to enable DEQ fine-tuning of DMs on inverse problems [19]. By refining a generative prior at the fixed-point of MAP estimation, MUPET enhances image restoration via posterior sampling while maintaining quality when sampling from the prior.

8.4.2 Taxonomy of reduction matrices for Graph Coarsening

Participants: Antonin Joly, Nicolas Keriven, Aline Roumy.

Graph coarsening aims to diminish the size of a graph to lighten its memory footprint, and has numerous applications in graph signal processing and machine learning. It is usually defined using a reduction matrix and a lifting matrix, which, respectively, allows to project a graph signal from the original graph to the coarsened one and back. This results in a loss of information measured by the so-called Restricted Spectral Approximation (RSA). Most coarsening frameworks impose a fixed relationship between the reduction and lifting matrices, generally as pseudo-inverses of each other, and seek to define a coarsening that minimizes the RSA. In [13], we remark that the roles of these two matrices are not entirely symmetric: indeed, putting constraints on the lifting matrix alone ensures the existence of important objects such as the coarsened graph's adjacency matrix or Laplacian. In light of this, in this paper, we introduce a more general notion of reduction matrix, that is not necessarily the pseudo-inverse of the lifting matrix. We establish a taxonomy of "admissible" families of reduction matrices, discuss the different properties that they must satisfy and whether they admit a closed-form description or not. We show that, for a fixed coarsening represented by a fixed lifting matrix, the RSA can be further reduced simply by modifying the reduction matrix. We explore different examples, including some based on a constrained optimization process of the RSA. Since this criterion has also been linked to the performance of Graph Neural Networks, we also illustrate the impact of this choices on different node classification tasks on coarsened graphs. This work was published at the NeurIPS conference.

8.4.3 Node Regression on Latent Position Random Graphs via Local Averaging

Participants: Nicolas Keriven.

Node regression consists in predicting the value of a graph label at a node, given observations at the other nodes. To gain some insight into the performance of various estimators for this task, in [7] we perform a theoretical study in a context where the graph is random. Specifically, we assume that the graph is generated by a Latent Position Model, where each node of the graph has a latent position, and the probability that two nodes are connected depend on the distance between the latent positions of the two nodes. In this context, we begin by studying the simplest possible estimator for graph regression, which consists in averaging the value of the label at all neighboring nodes. We show that in Latent Position Models this estimator tends to a Nadaraya Watson estimator in the latent space, and that its rate of convergence is in fact the same. One issue with this standard estimator is that it averages over a region consisting of all neighbors of a node, and that depending on the graph model this may be too much or too little. An alternative consists in first estimating the true distances between the latent positions, then injecting these estimated distances into a classical Nadaraya Watson estimator. This enables averaging in regions either smaller or larger than the typical graph neighborhood. We show that this method can achieve standard nonparametric rates in certain instances even when the graph neighborhood is too large or too small. This work was published in the Journal of Machine Learning Research (JMLR).

8.4.4 Backward Oversmoothing: why is it hard to train deep Graph Neural Networks?

Participants: Nicolas Keriven.

Oversmoothing has long been identified as a major limitation of Graph Neural Networks (GNNs): input node features are smoothed at each layer and converge to a non-informative representation, if the weights of the GNN are sufficiently bounded. This assumption is crucial: if, on the contrary, the weights are sufficiently large, then oversmoothing may not happen. Theoretically, GNN could thus learn to not oversmooth. However it does not really happen in practice, which prompts us to examine oversmoothing from an optimization point

of view. In the preprint [27], we analyze backward oversmoothing, that is, the notion that backpropagated errors used to compute gradients are also subject to oversmoothing from output to input. With non-linear activation functions, we outline the key role of the interaction between forward and backward smoothing. Moreover, we show that, due to backward oversmoothing, GNNs provably exhibit many spurious stationary points: as soon as the last layer is trained, the whole GNN is at a stationary point. As a result, we can exhibit regions where gradients are near-zero while the loss stays high. The proof relies on the fact that, unlike forward oversmoothing, backward errors are subjected to a linear oversmoothing even in the presence of non-linear activation function, such that the average of the output error plays a key role. Additionally, we show that this phenomenon is specific to deep GNNs, and exhibit counter-example Multi-Layer Perceptron. This paper is a step toward a more complete comprehension of the optimization landscape specific to GNNs.

9 Bilateral contracts and grants with industry

9.1 Bilateral contracts with industry

9.1.1 CIFRE contract with TyndallFx on Radiance fields representation for dynamic scene reconstruction

Participants: Christine Guillemot *(contact)*, Stephane Belemkoabga, Thomas Maugey.

- Title : Radiance fields representation for dynamic scene reconstruction
- Partners : TyndallFx (R. Mallart), Inria-Rennes.
- Funding : TyndallFx, ANRT.
- Period : Oct-2023-June. 2025

The goal of this project is to design novel methods for modeling and compact representation of radiance fields for scene reconstruction and view synthesis. The problems that are addressed are those of fast and efficient estimation of the camera pose parameters and of the 3D model of the sceen based on Gaussian splatting, and as as the one of tracking and modeling the deformation of the model due to the global camera motion and to the motion of the different objects in the scene.

9.1.2 CIFRE contract with MediaKind on Learned video downscaling for end-to-end Rate-Distortion optimization of video streaming system

Participants: Thomas Maugey *(contact)*, Esteban Pesnel, Aline Roumy.

- Title : Learned video downscaling for end-to-end Rate-Distortion optimization of video streaming system
- Partners : MediaKind, Inria-Rennes.
- Funding : MediaKind, ANRT.
- Period : November 2023-October 2026.

This CIFRE contract aims to optimize a streaming solution by addressing constraints related to distribution, standards, and deployment. The focus is on developing downscaling techniques that enhance the end-to-end streaming process, considering bitrate-distortion optimization. While the upscaling filter on client devices is fixed due to standardization, encoding and downscaling on the server side remain flexible, offering an opportunity for improvement within the streaming pipeline.

9.1.3 CIFRE contract with InterDigital on Hybrid conventional and deep learning-based video coding

Participants: Aline Roumy *(contact)*, Antoine Monier.

- Title : Hybrid conventional and deep learning-based video coding
- Partners : InterDigital, Inria-Rennes.
- Funding : InterDigital, ANRT.
- Period : Jan. 2025-Dec. 2028.

This CIFRE contract aims to improve conventional video codecs in terms of compression efficiency with the help of deep-learning and machine-learning based coding tools. The goal is to investigate the usage of deep-learning solutions for enhancing core video coding modules such as transform and residual (transform coefficients) coding, in-loop filtering, prediction. These new solutions should complement or replace existing coding tools or modes, such as the ones implemented in the VVC standard, or in the exploratory video coding model developed by the JVET standardization group named "Enhanced coding model" (ECM).

9.1.4 CIFRE contract with InterDigital on End-to-end energy-constrained video content delivery

Participants: Thomas Maugey *(contact)*, Emmanuel Sampaio.

- Title : End-to-end energy-constrained video content delivery
- Partners : InterDigital, Inria-Rennes.
- Funding : InterDigital, ANRT.
- Period : Jan. 2025-Dec. 2028.

The goal is to investigate new algorithms and video delivery frameworks to reduce the energy consumption footprint (and then the carbon footprint) of video content delivery. To reach this ambitious goal, several levers or strategies can be activated:

- Content pre-processing for reducing the encoding / transmission / decoding / rendering energy footprint. Assuming that the content is modified at the server side, this raises some important concerns: can we maintain the Quality of Experience? Can we guarantee an acceptance level? Do we need to provide side-information for making the process more efficient? If yes, is this overhead relevant for a commercial and viable operational deployment?
- Content post-processing for reducing the rendering energy footprint. Modifying the content at the client side raises the concern of the computational cost. A balance between energy gain and energy required to perform the post-processing operation has to be carefully considered.
- The delivery and consumption of video content are performed thanks to video streaming services. One of the key ingredients of such services relies on adaptive bitrate techniques aiming to deliver the highest QoE to the users given a bit rate constraint. We may want to go further by adding a new ingredient to the recipe, i.e., the energy consumption of such services. By considering the bit rate, the quality of experience and the energy footprint of the video, new energy-aware video streaming services could be envisioned

10 Partnerships and cooperations

10.1 European initiatives

10.1.1 Horizon Europe

ERC Starting Grant MALAGA: Reinventing the Theory of Machine Learning on Large Graphs

Participants: Nicolas Keriven *(PI)*, Hugo Jaquard, , Adarsh Jamadandi.

- Period: 2025 - 2030

In many scientific domains, graphs are the objects of choice to represent structured data: from molecules to social networks, power grids, the internet, and so on. The exploitation of graph data represents a major scientific and industrial challenge. Graph Machine Learning (Graph ML) is thus a fast-growing field, with so-called Graph Neural Networks (GNN) at the forefront.

However, in sharp contrast with traditional ML, the field of GML has somewhat jumped from early methods to deep learning, without the decades-long development of well-established notions to compare, analyze and improve algorithms. As a result, GNNs have limitations, both practical and theoretical, and it is not clear how to address them. Practical results may vary wildly depending on the architecture and datasets, with no guidelines on how to design reliable GNNs in each case. Overall, these are the symptoms of a major issue: Graph ML is somewhat lacking fundamental theory. The ambition of project MALAGA is to develop such a theory. Solving the crucial limitations of the current theory is highly challenging: fundamental mathematical tools in cannot analyze the learning capabilities of Graph ML methods in a unified way (e.g., graph nodes are not iid), existing *statistical graph models* do not faithfully represent the many characteristics of modern graph data (especially node features and their relationship with graph structure in homophilic and heterophilic graphs), and computational complexity may become problematic on large graphs. MALAGA will develop a radically new understanding of GML problems, and of the strengths and limitations of a large panel of algorithms.

10.1.2 H2020 projects

Participants: Christine Guillemot *(contact)*, Anil Ipek Atalay Appak, , Soheib Takhtardeshir, , Samuel Willingham.

- Title: Plenoptima: Plenoptic Imaging
- Duration: From January 1, 2021 to December 31, 2025
- Partners:
  - INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE (INRIA), France
  - MITTUNIVERSITETET (MIUN), Sweden
  - TECHNISCHE UNIVERSITAT BERLIN (TUB), Germany
  - TAMPEREEN KORKEAKOULUSAATIO SR (TAMPERE UNIVERSITY), Finland
  - "INSTITUTE OF OPTICAL MATERIALS AND TECHNOLOGIES ""ACADEMICIAN JORDAN MALINOWSKI"" - BULGARIAN ACADEMY OF SCIENCES" (IOMT), Bulgaria
- Inria contact: Christine Guillemot
- Coordinator: Tampere University (Finland, Atanas Gotchev)

Plenoptic Imaging aims at studying the phenomena of light field formation, propagation, sensing and perception along with the computational methods for extracting, processing and rendering the visual information.

The PLENOPTIMA ultimate project goal is to establish new cross-sectorial, international, multi-university sustainable doctoral degree programmes in the area of plenoptic imaging and to train the first fifteen future researchers and creative professionals within these programmes for the benefit of a variety of application sectors. PLENOPTIMA develops a cross-disciplinary approach to imaging, which includes the physics of light, new optical materials and sensing principles, signal processing methods, new computing architectures, and vision science modelling. With this aim, PLENOPTIMA joints five of strong research groups in nanophotonics, imaging and machine learning in Europe with twelve innovative companies, research institutes and a pre-competitive business ecosystem developing and marketing plenoptic imaging devices and services.

PLENOPTIMA advances the plenoptic imaging theory to set the foundations for developing future imaging systems that handle visual information in fundamentally new ways, augmenting the human perceptual, creative, and cognitive capabilities. More specifically, it develops 1) Full computational plenoptic imaging acquisition systems; 2) Pioneering models and methods for plenoptic data processing, with a focus on dimensionality reduction, compression, and inverse problems; 3) Efficient rendering and interactive visualization on immersive displays reproducing all physiological visual depth cues and enabling realistic interaction.

All ESRs are registered in Joint/Double degree doctoral programmes at academic institutions in Bulgaria, Finland, France, Germany and Sweden. The programmes will be made sustainable through a set of measures in accordance with the Salzburg II Recommendations of the European University Association.

10.2 National initiatives

10.2.1 PEPR MoleculArXiv. Targeted project 2: From Digital Data to Synthetic DNA

Participants: Aline Roumy *(contact)*, Sara Al Sayyed, , Thomas Maugey.

- Partners: I3S, LabSTIC, IMT-Atlantique, Irisa/Inria (GenScale and Compact team), IPMC, Eurecom.
- Funding: France 2030.
- Period: Sept. 2022 - Feb. 2032.

The PEPR MoleculArXiv aims to develop future data storage devices on molecular media, including DNA and artificial polymers. This involves not only parallelizing synthesis devices but also discovering new molecules and information technologies to accelerate the synthesis of storage media, their encoding and decoding, and exploring various molecular supports.

Within the targeted project "From Digital Data to Synthetic DNA," the objective is to make physical and logical storage efficient through custom-designed codes tailored to the physicochemical constraints of DNA writing and reading. This effort is conducted in collaboration with partners from other targeted projects, such as "Next-Generation DNA Synthesis" and "Synthetic Digital Polymers."

Several key challenges are addressed, including robustness to noise. Processes like synthesis, sequencing, storage, or manipulation of DNA can introduce errors that threaten the integrity of the stored data. These errors are non-classical compared to those encountered in wired and wireless communication channels and require specific handling. This issue is approached from the perspectives of both compression and error-correcting codes.

Another critical challenge is data access. A significant advantage of storing information on DNA, apart from its durability, is its extremely high density, enabling vast amounts of data to be stored compactly. Due to this high density, it is essential to facilitate rapid access to the required data items. New data representations are studied to enable fast random access to the data relying merely on biological and chemical processes.

10.2.2 PEPR IA. Project SHARP : Sharp Theoretical and Algorithmic Principles for frugal ML

Participants: Nicolas Keriven *(contact)*, Antonin Joly, , Caroline Mazini-Rodrigues.

- Partners: LIP, ENPC, IRISA, INRIA, CEA, LAMSADE, ISIR.
- Funding: France 2030.
- Period: 2023 - 2028

SHARP will address the major challenge of designing, analyzing and deploying a new generation of intrinsically frugal models (neural or not) able to achieve the versatility and performance of today's best models while requiring only a vanishing fraction of the resources currently needed. This will be achieved by the constitution of a strong task force able to cover an integrated pipeline, from theoretical foundations to flagship AI domains such as computer vision and natural language processing. With foundational advances towards stronger principles, smaller models, smaller datasets, SHARP will allow tomorrow's best AI systems to run on yesterday's devices, somewhat providing a cure against obsolescence.

10.2.3 ANR Young researcher grant: MAssive multimedia DAta collection REpurposing (MADARE)

Participants: Thomas Maugey *(contact)*, Tom Bordin.

- Funding: ANR (Agence Nationale de la Recherche)
- Period: pr. 2022 - Oct. 2025.

Compression algorithms are nowadays overwhelmed by the tsunami of visual data created everyday. Despite a growing efficiency, they are always constrained to minimize the compression error, computed in the pixel domain. The Data Repurposing framework, proposed in the MADARE project, will tear down this barrier, by allowing the compression algorithm to "reinvent" part of the data at the decoding phase, and thus saving a lot of bit-rate by not coding it. Concretely, a data collection is only encoded to a compact description that is used to guarantee that the regenerated content is semantically coherent with the initial one. In practice, it opens several research directions: how to organise the latent space (in which the coded descriptions lie) such that the information is efficiently and intelligibly represented? How to regenerate a synthesized content from this compact description (based for example on guided diffusion algorithms)? Finally, how to extend this idea to video? By revisiting the compression problem, the MADARE project aims gigantic compression ratios enabling, among other benefits, to reduce the impact of exploding data creation on the cloud servers' energy consumption.

10.2.4 Joint Project (Défi commun) Nisk.AI

Participants: Aline Roumy *(contact)*, Thomas Maugey, Antoine Monier, Christine Guillemot, Emmanuel Victor Barbosa Sampaio.

- Partners: Inria teams (Compact, Combo, Taran), InterDigital.
- Funding: Inria InterDigital.
- Period: Sept. 2022 - Feb. 2032.

Nisk.AI (2020-2026) is a joint project with InterDigital on Sustainable Neural Network video coding. Indeed, video distribution faces two major revolutions. The first one is due to the impact of AI technologies and in particular deep learning. New ways to represent images and video have been proposed by the scientific community and might impact how content is encoded, with very promising outputs in terms of coding

efficiency (e.g. the tradeoff between data-rate reduction and rendered perceived quality). The second revolution is the environmental impact of media consumption, and more generally of ICT (Information and Communication Technologies), on the global carbon footprint. This relates not only to the profusion of content and of its wide distribution, but also to how this content is processed and consumed, including users' behavior. The first revolution also has an impact on the second one due to the increased complexity of deep learning architectures compared to conventional coding schemas. The objective of this project is to address those challenges by proposing new deep-based video representation formats and coding schemes, taking into account efficiency, complexity and sustainability. Both 2D and immersive video will be considered.

10.3 Regional initiatives

10.3.1 CominLabs Colearn project: Coding for Learning

Participants: Aline Roumy *(contact)*, Rémi Piau, , Thomas Maugey.

- Partners: Inria-Rennes (Compact team); LabSTICC, IMT Atlantique, (team Code and SI3); IETR, INSA Rennes (Syscom team).
- Funding: Labex CominLabs.
- Period: Sept. 2021 - Dec. 2026.
- contact: Aline Roumy

The amount of data available online is growing so fast that it is essential to rely on advanced Machine Learning techniques so as to automatically analyze, sort, and organize the content uploaded by e.g. sensors or users. The conventional data transmission framework assumes that the data should be completely reconstructed, even with some distortions, by the server. Instead, this project aims to develop a novel communication framework in which the server may also apply a learning task over the coded data. The project will therefore develop an Information Theoretic analysis so as to understand the fundamental limits of such systems, and develop novel coding techniques allowing for both learning and data reconstruction from the coded data.

10.3.2 CominLabs VideoImpact project: Model the environmental cost of video delivery

Participants: Thomas Maugey *(contact)*, Natacha Lapeyroux, Robin Richard.

- Partners: MAGELLAN (IRISA/Inria), VAADER at IETR/INSA, ARENES University of Rennes, UCO Nantes, IMT Atlantique
- Funding: Labex CominLabs.
- Period: Sept. 2025 - Sep. 2027
- contact: Thomas Maugey

Recent studies forecast a global warming of 3.1°C in 2100 if the GHG emissions do not decrease. Hence, every part of our society must urgently aim sobriety, including the digital world, that is not intangible, contrary to popular belief. Video consumption takes a significant part among the emissions of the digital world and constitutes a representative example of unbounded and energy-consuming digital system. In that context, a crucial question to tackle is how to set limits to the deployment of a digital system, and for example to video delivery systems? This question is, by nature, lying at the crossroad of many fields (including human and social sciences). Interestingly, many initiatives have recently emerged at the regional level, e.g., the rapprochement between the GIS Marousin and video processing scientists of INSA and IRISA,

and set interesting perspectives of wide collaborative user experiments. In that context, the VideoImpact project proposes to answer the following questions: In order to set a sobriety policy, what should we limit in priority? the number of hours spent by a user watching videos? The TV screen size? The video resolutions? The deployment of more efficient digital infrastructure? The VideoImpact project aims at developing i) an environmental footprint model for the video delivery chain to identify the clear levers to sobriety, ii) a solid network of industrial and academic partners of the Rennes' neighborhood around the goal of reducing the environmental impact of video consumption and iii) to launch a concrete experimentation in collaboration with Human and Social scientists. The conclusions will be used in the context of further collaborations with Human and Social Scientists to set real user experiments to assess the feasibility and acceptance of such levers.

10.3.3 ARED VideoLimit project

Participants: Thomas Maugey *(contact)*, Robin Richard.

- Partners: MAGELLAN (IRISA/Inria)
- Funding: Labex CominLabs.
- Period: Sept. 2025 - Sep. 2028

In line with the project Cominlabs VideoImpact, the project Vlimit will specifically focus on the modeling of the energetic expense of the video transmission chain. More specifically the thesis funded by the project will focus on:

- model the energy spent over the whole video processing chain during different delivery scenarios, based on the state-of-the art analysis and experimental measurement campain.
- identify the high-energetic parts in this pipeline and some related levers that could be put in place to reduce their costs, based on a simulation tool for a « what-if » analysis.
- discuss with the Human and Social Siences researchers for setting the foundations of experimentations and inter-discplinary research directions, based on regular meetings and workshops with the active regional community

11 Dissemination

11.1 Promoting scientific activities

11.1.1 Scientific events: organisation

Member of the conference program committees

- Thomas Maugey was Area Chair for the EURASIP conference EUSIPCO 2025, Palermo, Italy
- Aline Roumy was a member of the technical program committee of the (Conference on Computer Vision and Pattern Recognition) CVPR 2025 workshop on New Trends in Image Restoration and Enhancement (NTIRE).
- Aline Roumy was a member of the technical program committee of the (International Conference on Computer Vision) ICCV 2025 workshop on Advances in Image Manipulation (AIM).
- Aline Roumy was a member of the technical program committee of the 2025 National Signal Processing workshop (colloque GRETSI).

Reviewer

- Thomas Maugey is reviewer for the following international conferences: EUSIPCO, ICIP, ICASSP, PCS
- Aline Roumy was a meta-reviewer for the 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) conference.
- Aline Roumy was a reviewer for the following international conferences: ICIP, ICASSP, ISIT

11.1.2 Journal

Member of the editorial boards

- Thomas Maugey is associate editor of the IEEE Signal Processing Letter.
- Aline Roumy is Senior Associate Editor of the IEEE Transactions on Image Processing.

Reviewer - reviewing activities

- Thomas Maugey is reviewer for IEEE Trans. on Image Processing and IEEE Signal Processing Letters

11.1.3 Invited talks

- Thomas Maugey gave a talk at L2S, Paris Saclay on "semantic compression: exploring ultra low bitrate" (January)
- Thomas Maugey gave a talk at the GdR meeting on "Sustainaibility and carbon footprint of the video transmission chain" on “Reducing environmental impact: from global modeling to behavioral change” (March)
- Thomas Maugey gave a talk at the VAADER semainar (IINSA IETR), on “Reducing environmental impact: from global modeling to behavioral change” (May)
- Thomas Maugey gave a talk at the Inria-InterDigital NEMO workshop on "semantic compression: exploring ultra low bitrate" (November)
- Aline Roumy gave a tutorial on “Information theory for image and video compression: fundamental results and recent challenges" MoleculArXiv Autumn School on DNA Data Storage, Nov. 2025.
- Aline Roumy gave a talk at the Inria-InterDigital NEMO workshop on "Image compression at JPEG: JPEG AI and JPEG DNA" (Nov. 2025)

11.1.4 Leadership within the scientific community

- Thomas Maugey is Vice-Chair of the EURASIP Technical Area Committee on Visual Information Processing
- Aline Roumy is a member of the IEEE Image, Video, and Multidimensional Signal Processing Technical Committee (IVMSP TC).
- Aline Roumy is a member of the Executive board of the National Research group in Image and Signal Processing (GRETSI).

11.1.5 Scientific expertise

- Christine Guillemot is member of the ERC PE7 Advanced grant panel.
- Christine Guillemot is member of the jury for the signal image vision PhD prize of the Club EEA, GdR IASIS and GRETSI.
- Aline Roumy has been a member of the jury for the recruitment of Inria Junior researcher (CRCN/ISFP) in Rennes, May 2025.
- Aline Roumy served as a member of Board of Examiners (Comité de sélection) for an assistant professor position (Maitres de Conférences) at Polytech Nantes University, May 2025.
- Aline Roumy has been a member of the committee for the French Academy of Sciences/Inria Awards, June 2025.
- Aline Roumy was a reviewer for the evaluation committee for the appointment of a professor, Telecom Paris, Sept. 2025.
- Aline Roumy served as a member of Board of Examiners (Comité de sélection) for a Professor position (Professeur des Universités) at CentraleSupélec, University, Oct. 2025.

11.1.6 Research administration

- Christine Guillemot is member of the ERC Cell of the DPE (Direction des Programmes Européens) of Inria.
- Aline Roumy is a member of the research commission and of the academic board of the University of Rennes 2, as Inria representative
- Aline Roumy is the co-director of the joint Inria/InterDigital project (défi) Nisk.AI

11.2 Teaching - Supervision - Juries - Educational and pedagogical outreach

- Thomas Maugey has given a course on Graph Image Processing, 10 hours, M2 SiVOS, Univ. of Rennes, France.
- Thomas Maugey has given a course on Ecological Transition and digital world, 6 hours, L3 SIF, ENS Rennes, France.
- Aline Roumy has given an Engineering degree course on the foundations of Image compression, 36 hours, University Rennes, ESIR, France.
- Aline Roumy has given an Engineering degree course on Image and Video compression, 10 hours, University Rennes, ESIR, France.

11.2.1 Supervision

- Thomas Maugey and Christine Guillemot were co-supervising the PhD thesis of Stéphane Belemkoabga in the context of the Cifre contract with TyndallFX.
- Thomas Maugey and Aline Roumy are co-supervising the PhD thesis of Esteban Pesnel in the context of the Cifre contract with Mediakind.
- Thomas Maugey and Aline Roumy were co-supervising the PhD thesis of Rémi Piau in the context of the Cominlabs project CoLearn.
- Thomas Maugey and Aline Roumy are co-supervising the PhD thesis of Sara Al Sayyed in the context of the PEPR project MoleculArxiv.

- Thomas Maugey is co-supervising the PhD thesis of Emmanuel Sampaio in the context of the Cifre contract with InterDigital.
- Thomas Maugey was supervising the PhD thesis of Tom Bordin in the context of the ANR project MADARE
- Thomas Maugey is co-supervising the PhD thesis of Robin Richard in the context of the Bretagne ARED contract.
- Christine Guillemot is co-supervising Soheib Takhtardeshir together with Marten Sjostrom from MidSweden University in the context of the Plenoptima Marie Curie project
- Christine Guillemot is co-supervising Samuel Willigham together with Marten Sjostrom from MidSweden University in the context of the Plenoptima Marie Curie project
- Christine Guillemot is co-supervising Ipek Anil Atalay Appak together with Humeyra Caglayan from Tampere University in the context of the Plenoptima Marie Curie project
- Christine Guillemot is co-supervising Leo-Paul Huar together with Pierre Hellier in the context of a Cifre contract with InterDigital.
- Nicolas Keriven and Aline Roumy are co-supervising the PhD thesis of Antonin Joly in the context of the PEPR SHARP
- Nicolas Keriven and Aline Roumy are co-supervising the PhD thesis of Adarsh Jamadandi in the context of the ERC MALAGA
- Aline Roumy is co-supervising the PhD thesis of Antoine Monier with Pierre Hellier in the context of the joint Inria/InterDigital research project (defi commun) Nisk.AI.

11.2.2 Juries

- Christine Guillemot was member, as chair, of the PhD jury of Shubhendu JENA of the University of Rennes, June 2025.
- Christine Guillemot was member, as rapporteur, of the PhD jury of Aytaç Özkan at the Technical University of Berlin, Dec. 2025.
- Thomas Maugey was member, as examiner, of the PhD jury of Goluck KONUKO at the Paris-Saclay University, Jan. 2025.
- Thomas Maugey was member, as President, of the PhD jury of Sébastien DAM at the Rennes University, Oct. 2025.
- Thomas Maugey was member, as rapporteur, of the PhD jury of Gabriele SPADARO at TELECOM Paris Institut Polytechnique, Dec. 2025.
- Aline Roumy was member of the PhD committee of Corentin Presvôts, Paris-Saclay University, Jan. 2025, as a chair.
- Aline Roumy was member of the PhD committee of Rodrigo Borba Pinheiro, Paris-Saclay University, Jan. 2025, as a chair.
- Aline Roumy was member of the PhD committee of Jeremy Jaspar, Sorbonne Paris-Nord University, March. 2025, as a reviewer.
- Aline Roumy was member of the PhD committee of Pierre-Alain.Afro, Grenoble Alpes University, April. 2025, as a reviewer.
- Aline Roumy was member of the PhD committee of Maxime Ossonce, Paris-Saclay University, Dec. 2025, as an examiner.

11.2.3 Internal or external Inria responsibilities

- Aline Roumy is a member of the Gender Equality committee of Inria-Rennes and Irisa, responsible for the working group on career interruptions and support.
- Aline Roumy is a member of the mentorship program as a mentor.
- Thomas Maugey is a member of the *Formation Spécialisée de Site*, responsible of the security at work
- Thomas Maugey is a member of the SEnS group, animating the reflexion on our research goals and impacts at the level of the laboratory.

11.3 Popularization

11.3.1 Specific official responsibilities in science outreach structures

- Thomas Maugey is Scientific mediation officer in the scientific mediation team of Inria centre at Rennes Universiy.

11.3.2 Productions (articles, videos, podcasts, serious games, ...)

- Thomas Maugey is the co-designer and co-supervisor of the project *Ma thèse une sacré histoire*

11.3.3 Participation in Live events

- Thomas Maugey attended the "scientific mediation days of Inria" at the Ministère de l'enseignement supérieur et de la recherche, and did a presentation on the *ma thèse une sacré histoire* project.

12 Scientific production

12.1 Major publications

[1] T. Bordin and T. Maugey. 'Linearly transformed color guide for low-bitrate diffusion based image compression'. In: *IEEE Transactions on Image Processing* (18th Dec. 2024), p. 15. URL: https://inria.hal.science/hal-04882103. In press.

[2] N. Charpenay, M. Le Treust and A. Roumy. 'Side Information Design in Zero-Error Coding for Computing'. In: *Entropy* 26.4 (16th Apr. 2024), pp. 1–18. DOI: 10.3390/e26040338. URL: https://hal.science/hal-04679240.

[3] A. Joly and N. Keriven. 'Graph Coarsening with Message-Passing Guarantees'. In: *Advances in Neural Information Processing Systems (NeurIPS)*. Advances in Neural Information Processing Systems (NeurIPS). Vancouver, Canada, 2024. URL: https://hal.science/hal-04617519.

[4] M. Le Pendu and C. Guillemot. 'Preconditioned Plug-and-Play ADMM with Locally Adjustable Denoiser for Image Restoration Mikael'. In: *SIAM Journal on Imaging Sciences* (Nov. 2022), pp. 1–30. URL: https://hal.science/hal-03857826.

12.2 Publications of the year

International journals

[5] N. Charpenay, M. L. Treust and A. Roumy. 'On the Additivity of Optimal Rates for Independent Zero-Error Source and Channel Problems'. In: *IEEE Transactions on Information Theory* (2025), pp. 1–1. DOI: 10.1109/TIT.2025.3639292. URL: https://inria.hal.science/hal-05401169 (cit. on pp. 14, 16).

[6] D. R. Freitas, I. Tabus and C. Guillemot. 'Visibility-Based Geometry Pruning of Neural Plenoptic Scene Representations'. In: *IEEE Transactions on Multimedia* (2025), pp. 1–17. DOI: 10.1109/tmm.2025.3618548. URL: https://hal.science/hal-05101627. In press.

[7] M. Gjorgjevski, N. Keriven, S. Barthelme and Y. D. Castro. 'Node Regression on Latent Position Random Graphs via Local Averaging'. In: *Journal of Machine Learning Research* (2025). URL: https://hal.science/hal-04795829. In press (cit. on p. 22).

[8] S. Takhtardeshir, R. Olsson, C. Guillemot and M. Sjöström. 'DUALF-D: Disentangled dual-hyperprior approach for light field image compression'. In: *Signal Processing: Image Communication* 140 (Jan. 2026), p. 117436. DOI: 10.1016/j.image.2025.117436. URL: https://hal.science/hal-05372515 (cit. on p. 16).

[9] S. Takhtardeshir, R. Olsson, C. Guillemot and M. Sjöström. 'Efficient and Fast Light Field Compression via VAE-Based Spatial and Angular Disentanglement'. In: *IEEE Access* 13 (2025), pp. 18594–18607. DOI: 10.1109/ACCESS.2025.3532608. URL: https://hal.science/hal-05233427 (cit. on p. 16).

International peer-reviewed conferences

[10] S. Al Sayyed, A. Roumy and T. Maugey. 'Efficient Constraining of Transcoding in DNA-Based Image Storage'. In: ICIP 2025 - IEEE International Conference on Image Processing. Anchorage (AK), USA, United States: IEEE, 2025, pp. 1–6. URL: https://hal.science/hal-05300068 (cit. on p. 17).

[11] S. Al Sayyed, A. Roumy, T. Maugey, N. Lobato-Dauzier and A. J. Genot. 'Compact image representation for content-based image retrieval in DNA data storage'. In: *Picture Coding Symposium (PCS)*. PCS 2025 - Picture Coding Symposium. Aachen, Germany, 2025, pp. 1–5. URL: https://hal.science/hal-05322329 (cit. on p. 17).

[12] A. Jamadandi, J. Xu, A. Dziedzic and F. Boenisch. 'Memorization in Graph Neural Networks'. In: NeurIPS 2025 - 39th Conference on Neural Information Processing Systems. San Diego (CA), United States, 2025. URL: https://inria.hal.science/hal-05455794.

[13] A. Joly, N. Keriven and A. Roumy. 'Taxonomy of reduction matrices for Graph Coarsening'. In: *Advances in Neural Information Processing Systems (NeurIPS)*. NeurIPS 2025 - 39th Annual Conference on Neural Information Processing Systems. San diego, CA, United States, 2nd Dec. 2025. URL: https://inria.hal.science/hal-05248172 (cit. on pp. 14, 22).

[14] R. Kaafarani, J. L. Tanou, M. Ropert, T. Maugey and A. Roumy. 'Joint multi-profile video coding for cost optimization of standard ABR streaming'. In: *ACM Mile-High Video Conference (MHV 2025)*. MHV 2025 - 4th ACM Mile-High Video Conference. Denver (CO), United States, 18th Feb. 2025, pp. 33–39. DOI: 10.1145/3715675.3715795. URL: https://inria.hal.science/hal-05291651.

[15] E. Pesnel, J. Le Tanou, M. Ropert, T. Maugey and A. Roumy. 'SCALED : Surrogate-gradient for Codec-Aware Learning of Downsampling in ABR Streaming'. In: *Picture Coding Symposium (PCS)*. PCS 2025 - Picture Coding Symposium. Aachen (Aix la Chapelle), Germany, 8th Dec. 2025. URL: https://hal.science/hal-05301489 (cit. on p. 18).

[16] D. Stéphane Belemkoabga, T. Maugey and C. Guillemot. 'GS-Morph: Dynamic Novel View Synthesis via UDF-ARAP Gaussian Splat Morphing'. In: CVMP 2025 - ACM SIGGRAPH European Conference on Visual Media Production. London, United Kingdom: ACM, 2025, pp. 1–10. DOI: 10.1145/3756863.3769701. URL: https://inria.hal.science/hal-05311279 (cit. on pp. 14, 20).

[17] S. Takhtardeshir, R. Olsson, C. Guillemot and M. Sjöström. 'DUALF-C: Disentangled Light Field Compression with Entropy-Aware Bitstream Generation'. In: VCIP 2025 - International Conference on Visual Communications and Image Processing. Klagenfurt, Austria, 2025, pp. 1–5. URL: https://hal.science/hal-05327806 (cit. on p. 16).

[18] P. Wawerek-López, N. Mahmoudian Bidgoli, P. Frossard, A. Kaup and T. Maugey. 'OSLO-IC: On-the-Sphere Learned Omnidirectional Image Compression with Attention Modules and Spatial Context'. In: ICASSP 2025 - IEEE International Conference on Acoustics, Speech, and Signal Processing. Hyderabad, India, 2025, pp. 1–4. URL: https://inria.hal.science/hal-05442204 (cit. on p. 18).

[19] S. Willingham, M. Sjöström and C. Guillemot. 'Maximum A Posteriori Training of Diffusion Models for Image Restoration'. In: *EUSIPCO 2025 – 33rd European Signal Processing Conference*. EUSIPCO 2025 - 33rd European Signal Processing Conference. Isola delle Femmine, Italy, 2025, pp. 1–5. URL: https://inria.hal.science/hal-05280965 (cit. on p. 21).

National peer-reviewed Conferences

[20] M. Gjorgjevski, N. Keriven, S. Barthelme and Y. D. Castro. 'Graphical Kernel Ridge Regression in Latent Position Models'. In: *GRETSI 2025 - 30e Colloque Francophone sur le Traitement du Signal et des Images*. GRETSI 2025 - 30e Colloque Francophone sur le Traitement du Signal et des Images. Strasbourg, France, 2025, pp. 1–4. URL: https://hal.science/hal-05385113.

[21] A. Joly, N. Keriven and A. Roumy. 'Matrices de réduction et de reconstruction pour la réduction de graphes'. In: GRETSI 2025 - 30ème Colloque Francophone de Traitement du Signal et des Images. Strasbourg, France, 2025, pp. 1–4. URL: https://hal.science/hal-05248082.

[22] T. Maugey, D. Menard, A.-C. Orgerie and R. Richard. 'Diffusion vidéo : dépasser l'objectif d'efficacité et viser une sobriété choisie'. In: *GRETSI - XXXe Colloque Francophone de Traitement du Signal et des Images*. GRETSI 2025 - XXXe Colloque Francophone de Traitement du Signal et des Images. Strasbourg, France, 2025, pp. 1–4. URL: https://hal.science/hal-05223926 (cit. on p. 20).

Doctoral dissertations and habilitation theses

[23] T. Bordin. 'Image semantic compression at extremely low bitrates'. Université de Rennes 1, 4th Dec. 2025. URL: https://hal.science/tel-05454814.

[24] D. R. d. C. Freitas. 'Sensing and reconstruction of plenoptic point clouds'. Université de Tampere (Finlande), 21st May 2025. URL: https://theses.hal.science/tel-05253921.

[25] R. Piau. 'Learning from entropy-encoded data'. Université de Rennes, 26th June 2025. URL: https://hal.science/tel-05455203 (cit. on p. 16).

Reports & preprints

[26] H. Jaquard and N. Keriven. *Statistical Consistency of Discrete-to-Continuous Limits of Determinantal Point Processes*. 2026. URL: https://hal.science/hal-05448944.

[27] N. Keriven. *Backward Oversmoothing: why is it hard to train deep Graph Neural Networks?* 2025. URL: https://hal.science/hal-05238439 (cit. on p. 23).

[28] C. Mazini Rodrigues, N. Keriven and T. Maugey. *Compressing image encoders via latent distillation*. 2026. URL: https://hal.science/hal-05448409 (cit. on p. 19).

12.3 Cited publications

[29] J. J. Ahmad, H. A. Khan and S. A. Khayam. 'Energy efficient video compression for wireless sensor networks'. In: *2009 43rd Annual Conference on Information Sciences and Systems*. IEEE. 2009, pp. 629–634 (cit. on p. 10).

[30] A. Atalay Appak, E. Sahin, C. Guillemot and H. Caglayan. 'Learning flat optics for extended depth of field microscopy imaging'. In: *Nanophotonics* (2023) (cit. on p. 11).

[31] S. Bai, J. Z. Kolter and V. Koltun. 'dDeep Equilibrium Models'. In: *NEURIPS*. 2019 (cit. on p. 12).

[32] Y. Blau and T. Michaeli. 'The perception-distortion tradeoff'. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 6228–6237 (cit. on p. 9).

[33] M. M. Bronstein, J. Bruna, T. Cohen and P. Veličković. 'Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges'. In: (2021) (cit. on p. 12).

[34] B. Bross, Y.-K. Wang, Y. Ye, S. Liu, J. Chen, G. J. Sullivan and J.-R. Ohm. 'Overview of the versatile video coding (VVC) standard and its applications'. In: *IEEE Transactions on Circuits and Systems for Video Technology* 31.10 (2021), pp. 3736–3764 (cit. on p. 10).

[35] L. Ceze, J. Nivala and K. Strauss. ‘Molecular digital data storage using DNA’. In: *Nature Reviews Genetics* 20.8 (2019), pp. 456–466 (cit. on p. 8).

[36] L. D. Chamain, F. Racapé, J. Bégaint, A. Pushparaja and S. Feltman. ‘End-to-End optimized image compression for machines, a study’. In: *2021 Data Compression Conference (DCC)*. ISSN: 2375-0359. Mar. 2021, pp. 163–172. DOI: 10.1109/DCC50243.2021.00024 (cit. on p. 8).

[37] H. Choi and I. V. Bajić. ‘Scalable Image Coding for Humans and Machines’. In: *IEEE Transactions on Image Processing* 31 (2022). Conference Name: IEEE Transactions on Image Processing, pp. 2739–2754. DOI: 10.1109/TIP.2022.3160602 (cit. on p. 8).

[38] Copernicus. *Access to data*. en. https://www.copernicus.eu/en/access-data. Aug. 2023. URL: https://www.copernicus.eu/en/access-data (cit. on p. 5).

[39] G. Cormode, M. Garofalakis, P. J. Haas and C. Jermaine. ‘Synopses for Massive Data: Samples, Histograms, Wavelets, Sketches’. In: *Foundations and Trends in Databases* 4 (2011), pp. 1–294 (cit. on p. 12).

[40] Domo. *Data Never Sleeps 10.0*. en. https://www.domo.com/data-never-sleeps. June 2022 (cit. on pp. 5, 8).

[41] L. Duan, J. Liu, W. Yang, T. Huang and W. Gao. ‘Video coding for machines: A paradigm of collaborative compression and intelligent analytics’. In: *IEEE Transactions on Image Processing* 29 (2020), pp. 8680–8695 (cit. on p. 8).

[42] Ericsson. *Mobility Report*. en. https://www.ericsson.com/en/reports-and-papers/mobility-report. June 2023. URL: https://www.ericsson.com/en/reports-and-papers/mobility-report (cit. on p. 5).

[43] Y. Feng and A. Varshney. ‘Signet: Efficient neural representation for light fields’. In: *IEEE/CVF International Conference on Computer Vision (ICCV)*. 2021 (cit. on p. 11).

[44] W. Gao, S. Liu, X. Xu, M. Rafie, Y. Zhang and I. Curcio. ‘Recent standard development activities on video coding for machines’. In: *arXiv preprint arXiv:2105.12653* (2021) (cit. on p. 8).

[45] A. Kipnis, S. Rini and A. J. Goldsmith. ‘The Rate-Distortion Risk in Estimation From Compressed Data’. In: *IEEE Transactions on Information Theory* 67.5 (May 2021). Conference Name: IEEE Transactions on Information Theory, pp. 2910–2924. DOI: 10.1109/TIT.2021.3068981 (cit. on p. 8).

[46] B. Le Bon, M. Le Pendu and C. Guillemot. ‘Stochastic Unrolled Proximal Point Algorithm for linear image inverse problems’. In: *EUSIPCO 2023 - 31st European Signal Processing Conference*. Helsinki, Finland, 2023 (cit. on p. 12).

[47] G. Le Guludec and C. Guillemot. ‘Joint NeuraL Representation For Multiple Light Fields’. In: *ICASSP 2023 - IEEE Internal Conference on Acoustics, Speech and Signal Processing*. Ed. by IEEE. Rhodes, Greece: IEEE, June 2023, pp. 1–5. URL: https://hal.science/hal-04054325 (cit. on p. 11).

[48] J. Leskovec and C. Faloutsos. ‘Sampling from large graphs’. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 2006 (2006), pp. 631–636 (cit. on p. 13).

[49] D. Lim, H. Maron, M. T. Law, J. Lorraine and J. Lucas. ‘Graph Metanetworks for Processing Diverse Neural Architectures’. In: *International Conference on Learning Representations (ICLR)*. 2024 (cit. on pp. 12, 13).

[50] A. Loukas. ‘Graph reduction with spectral and cut guarantees’. In: *Journal of Machine Learning Research* 20 (2019), pp. 1–42 (cit. on p. 13).

[51] ‘Low-complexity video compression for wireless sensor networks’. In: *2003 International Conference on Multimedia and Expo. ICME’03. Proceedings (Cat. No. 03TH8698)*. Vol. 3. IEEE. 2003, pp. III–585 (cit. on p. 10).

[52] Z. Lu, Y. Liu and M. e. a. Jin. ‘Virtual-scanning light-field microscopy for robust snapshot high-resolution volumetric imaging’. In: *Nat Methods* (2023) (cit. on p. 11).

[53] V. Masson-Delmotte, P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P. R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock et al. 'Global warming of 1.5 C'. In: *An IPCC Special Report on the impacts of global warming of* 1.5 (2018), pp. 43–50 (cit. on p. 9).

[54] F. Mentzer, E. Agustsson, J. Ballé, D. Minnen, N. Johnston and G. Toderici. 'Neural video compression using gans for detail synthesis and propagation'. In: *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXVI*. Springer. 2022, pp. 562–578 (cit. on p. 10).

[55] F. Mentzer, G. D. Toderici, M. Tschannen and E. Agustsson. 'High-fidelity generative image compression'. In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 11913–11924 (cit. on p. 10).

[56] B. e. a. Mildenhall. 'Nerf: Representing scenes as neural radiance fields for view synthesis'. In: *ECCV*. 2020 (cit. on p. 11).

[57] S. Mohan. 'Robust and Interpretable Denoising Via Deep Learning'. PhD thesis. New York University, 2022 (cit. on p. 10).

[58] S. Mohan, Z. Kadkhodaie, E. P. Simoncelli and C. Fernandez-Granda. 'Robust and interpretable blind image denoising via bias-free convolutional neural networks'. In: *arXiv preprint arXiv:1906.05478* (2019) (cit. on p. 10).

[59] Z. Pan, H. Qin, X. Yi, Y. Zheng and A. Khan. 'Low complexity versatile video coding for traffic surveillance system'. In: *International Journal of Sensor Networks* 30.2 (2019), pp. 116–125 (cit. on p. 10).

[60] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu and M. Chen. 'Hierarchical text-conditional image generation with clip latents'. In: *arXiv preprint arXiv:2204.06125* (2022) (cit. on p. 9).

[61] D. R.-J. G.-J. Rydning, J. Reinsel and J. Gantz. 'The digitization of the world from edge to core'. In: *Framingham: International Data Corporation* 16 (2018), pp. 1–28 (cit. on p. 8).

[62] Sandvine. *Global Internet Phenomena Report*. https://www.sandvine.com/global-internet-phenomena-report-2023. 2023. URL: https://www.sandvine.com/global-internet-phenomena-report-2023 (cit. on p. 5).

[63] R. Sims, S. A. Rehman and M. O. L. et al. 'Single molecule light field microscopy'. In: *Optica* (2020) (cit. on p. 11).

[64] R. Torfason, F. Mentzer, E. Agustsson, M. Tschannen, R. Timofte and L. Van Gool. 'Towards Image Understanding from Deep Compression without Decoding'. In: *Int. Conf. on Learning Representations (ICLR)*. 2018. URL: http://arxiv.org/abs/1803.06131 (visited on 10/02/2020) (cit. on p. 8).

[65] J. P. Vizcaíno, F. Saltarin, Y. Belyaev, R. Lyck, T. Lasser and P. Favaro. 'Learning to Reconstruct Confocal Microscopy Stacks From Single Light Field Images'. In: *IEEE Transactions on Computational Imaging* 7 (2021), pp. 775–788. DOI: 10.1109/TCI.2021.3097611 (cit. on p. 11).

[66] Y. Yang, S. Mandt and L. Theis. 'An Introduction to Neural Data Compression'. In: *Foundations and Trends in Computer Graphics and Vision* 15.2 (2023), pp. 113–200 (cit. on p. 12).

[67] R. York and J. A. McGee. 'Understanding the Jevons paradox'. In: *Environmental Sociology* 2.1 (2016), pp. 77–87 (cit. on p. 10).

[68] V. Zhirnov, R. M. Zadegan, G. S. Sandhu, G. M. Church and W. L. Hughes. 'Nucleic acid memory'. In: *Nature materials* 15.4 (2016), pp. 366–370 (cit. on p. 8).

[69] Z. Zhongming, L. Linong, Y. Xiaona, Z. Wangqiang, L. Wei et al. 'AR6 synthesis report: Climate change 2022'. In: (2022) (cit. on p. 9).