INTERNSHIP OFFER
2024VDO-01
Energy-aware adaptative streaming of the video chain

LAB	VDO	Group	Meta Video	PROJECT	Energy Aware Media
Main Supervisor	Franck AUMONT				
Additional supervisors	Olivier LE MEUR				

Subject Proposal	
	Description: In the context of climate change, reducing the carbon emission generated by video streaming is a strong requirement. Video content represents 82% of the global internet traffic (Cisco's report [1]) and around 3.7 billion people world-wide will use video on demand services in 2027. State of the art of the adaptative streaming technologies allows to deliver video with the highest possible visual quality while influenced by various factors such as network bandwidth, display resolution and viewing conditions but not taking into account the reduction objective of the end-user device energy consumption. In this internship, the objective is to influence the adaptative streaming system to have an impact on the energy consumption of the end-user device without compromising the user experience while rendering the video. This work shall impact the entire video streaming chain (i.e., from the headend up to the display) by influencing the various factors (e.g., bitrate ladder, paper [2]) of the different components (i.e., encoder, packager, …) of the chain. The goal will be to review existing solutions, then investigate how to adapt the most promising adaptative streaming system to introduce an energy reduction factor, possibly developing a deep-learning or reinforcement learning based network. It will also be required to validate the principle in the entire video streaming chain. A benchmark will be realized to estimate the energy consumption gain and check that the balance between a high-quality user experience and energy reduction is possible. [1] U. Cisco, "Cisco annual internet report (2018–2023) white paper," *Cisco: San Jose, CA, USA*, 2020. [2] Ahmed Telili, Wassim Hamidouche, Sid Ahmed Fezza and Luce Morin, "Benchmarking Learning-based Bitrate Ladder Prediction Methods for Adaptive Video Streaming" Skills: Adaptative streaming (e.g., HLS, DASH), video coding and processing. Deep learning, reinforcement learning. C/C++, Python. PyTorch or an equivalent deep learning framework would be appreciated. Keywords: energy consumption, energy reduction, energy-aware adaptative streaming, vision quality metrics, machine learning (deep and reinforcement learning).

Expected outcomes		
Patents	Yes	As an innovative research axis, this internship will be expected to generate patents and publications. Software that demonstrates the feasibility and performance of the developed technique would also be expected.
Publications	Yes	
Demos	Yes	
Software modules	Yes	

INTERNSHIP OFFER
2024VDO-02

Display Energy Reduction through Energy Aware Color Replacement in Videos

LAB	VDO	Group	Meta Video	PROJECT	Energy Aware Media
Main Supervisor	Claire-Hélène Demarty				
Additional supervisors	Laurent Blondé				

Subject Proposal	
	Description: In the context of climate change, reducing the energy consumption of electronic devices is a requirement. As displays represent a large part of the consumption of such devices, they are a choice target to reduce our energy footprint. This internship will investigate different **chromaticity-based** methods, derived from properties of the visual perception, to reduce the energy needed to display **videos onscreen**. This internship will build upon an already existing technology in the team dedicated to the processing of still images to extend it to videos. Besides the development of the methods through classic computer vision algorithms, a deep-learning based network could be developed to mimic the most promising investigated technique. Standardization is as well a goal for the technology to become largely used and have an impact on worldwide displays consumption reduction. Skills: computer vision/ signal/ color/ image/ video processing. Background in display physics and deep learning would be appreciated. Python. PyTorch or an equivalent deep learning framework. Keywords: color, visual perception, energy reduction, energy aware images, computer vision, machine learning (deep learning).

Expected outcomes		
Patents	Yes	As an innovative research axis, this internship will be expected to generate patents and publications. A software module that demonstrates the feasibility and performance of the developed technique would also be expected.
Publications	Yes	
Demos	Yes	
Software modules	Yes	

INTERNSHIP OFFER
2024VDO-03

Face Analysis in Video Conference Using Synthetic Data

LAB	VDO	Group	MetaVideo	PROJECT	Interactive Media
Main Supervisor	Quentin Avril				
Additional supervisors	Philippe-Henri Gosselin				

Proposal	
	**Description**: Interdigital Immersive Video lab is working on semantic-based technology for videoconferences. In this context human facial 3D data is key to estimate semantic features such as facial expressions correctly and precisely, lighting conditions (eg. light positions or specular reflections), background and hair segmentation etc. In 2023, Interdigital built its own synthetic data generation framework enabling to create facial images. The goal of the internship will be to: 1. Set-up specific synthetic face datasets 2. Train neural networks and validate their higher accuracy The intern will be included in the Video lab in Rennes, France within the Interactive Media project. All the 3D assets (3D meshes, textures, facial expression) are already available. Professional tools such as Autodesk Maya or SideFX Houdini will also be provided in complement to a geometry processing framework. **Skills**: computer graphics, 3D modeling, rendering, Maya/Houdini/Blender **Keywords**: computer graphics, synthetic data, machine learning

Expected outcomes		
Patents	Yes	Face analysis for avatar animation or semantic streaming is under discussions within MPEG-SD and MPEG-3DGH
Publications	Yes	If the results are better than SoA

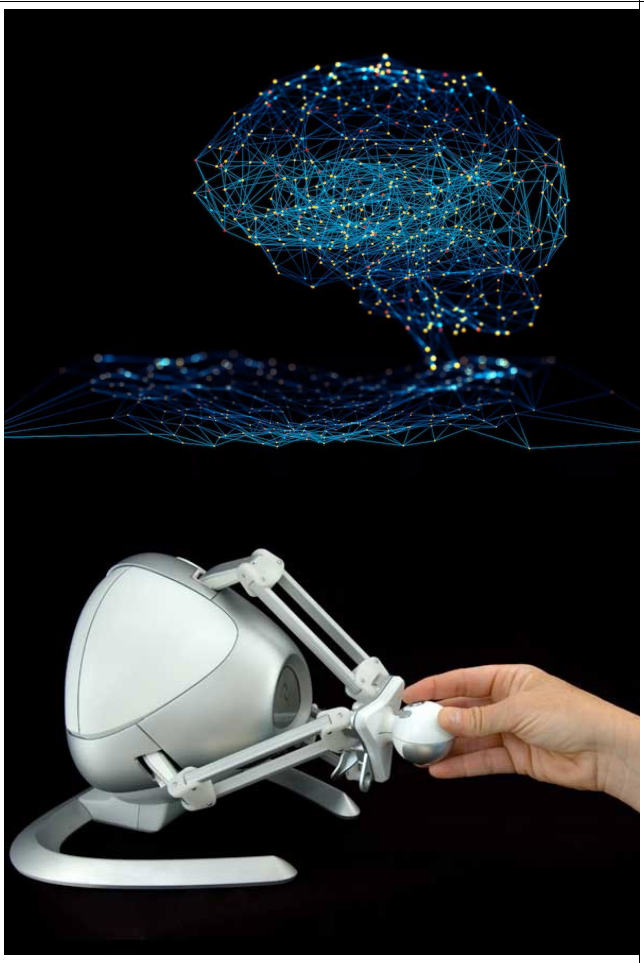
Demos	Yes	This will be included in the avatar demo
Software modules	Yes	It will be part of the avatar platform

INTERNSHIP OFFER
2024VDO-04

Implicit Neural Representation of Haptic data

LAB	VDO	Group	MetaVideo	PROJECT	Interactive Media
Main Supervisor	Gurvan Lécuyer				
Additional supervisors	Quentin Galvane, Philippe Guillotel				

Proposal



Description:
The intern will be joining the haptic group within the Meta Video team at Interdigital (Rennes). Implicit Neural Representation (INR) is an application of neural networks that aimed at parameterizing a continuous differentiable signal within the weights of the model. This can be used to encode 3D scenes from 2D images for example [1]. INRs are growing in popularity and are now used for various tasks not limited to vision. In the context of the internship, the goal is to explore the feasibility of applying INR to encode and render haptic data.

The work will consist in:
(1) study the state of the art on Implicit Neural Representation [1,2] and Haptics [3,4],
(2) propose new solutions to leverage Implicit Neural Representation for the encoding and rendering of Haptic data,
(3) implement, test, and evaluate the proposed solutions for a potential publication.

[1] Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2021). Nerf: Representing scenes as neural radiance fields for view synthesis. Com. of the ACM, 65(1), 99-106.
[2] Ben-Shabat, Y., Koneputugodage, C. H., & Gould, S. (2022). Digs: Divergence guided shape implicit neural representation for unoriented point clouds. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 19323-19332).
[3] Kim, K., & Park, J. (2019). Voxel-based haptic rendering using adaptive sampling of a local distance map. Journal of Computing Science and Engineering, 13(2), 66-77.

	[4] Rekik, Yosra & Vezzoli, Eric & Frédéric, Giraud & Grisoni, Laurent. (2017). Localized Haptic Texture: A Rendering Technique Based on Taxels for High Density Tactile Feedback. 10.1145/3025453.3026010. Skills: C++, Python, Deep Learning framework (Tensorflow or Pytorch), English Keywords: Haptics, Implicit Neural Representation

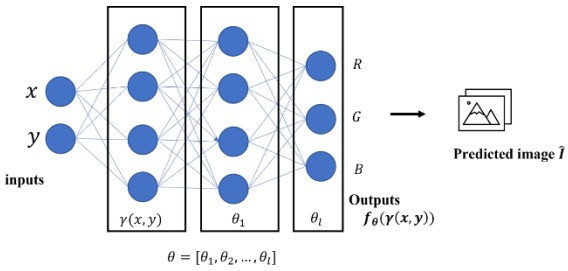
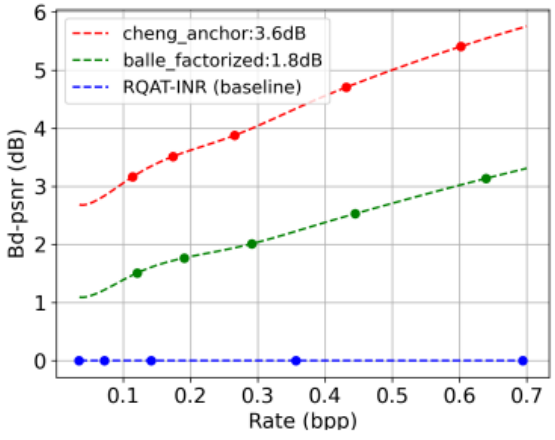
Expected outcomes		
Patents	Yes	Proposed solutions could be patented
Publications	Yes	The objective is to publish the finding of this internship
Demos	No	
Software modules	Yes	Software Modules should be developed as part of the research effort

INTERNSHIP OFFER
2024VDO-06

INR video compression

Dictionary learning and update methods for random access and low-delay modes

LAB	VDO	Group	MetaVideo	PROJECT	SympAI
Main Supervisor	Bharath Damodaran				
Additional supervisors	Pierre Hellier				

Proposal	
 	Description: Implicit neural representations (INR) have recently emerged as a new tool to model and represent various types of signals, including images [1,2,3] and videos [4]. In particular for image and video compression, these implicit representations can pave the way towards a radically new paradigm: once overfitted on an image to be encoded, transmitting the image amounts to transmitting the weights of the network, and the decoding is a straightforward evaluation of the neural network for each decoded pixel. Although promising, these methods [2,3,4] have not yet reached the performance of current state-of-the-art traditional and learning based approaches. This internship proposes to study further to improve these implicit representations for image and video compression. More specifically at InterDigital, we develop dictionary-based methods to tackle the spatial and temporal redundancy of the signal. The Internship proposes to improve the dictionary learning approach for video compression, by exploring optimized techniques for dictionary learning and dictionary update. In particular, the two classical compression modes will be investigated: random access (where all frames are known by the encoder), and low delay (where the encoder only knows the previous frames). In these two different contexts, the dictionary learning, transmission and update will be different, which is the main focus of the internship proposal.

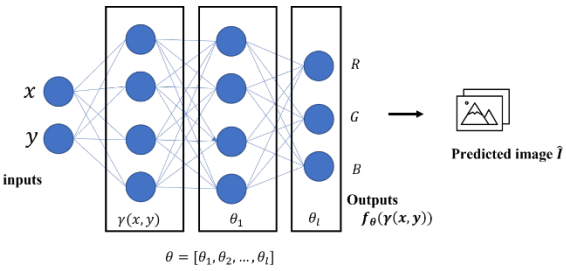
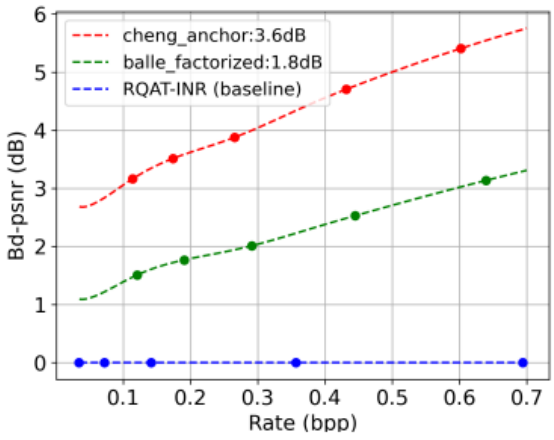
	References [1] Sitzmann, V., Martel, J., Bergman, A., Lindell, D., & Wetzstein, G. (2020). Implicit neural representations with periodic activation functions. Advances in Neural Information Processing Systems, 33, 7462-7473. [2] Dupont, E., Loya, H., Alizadeh, M., Goliński, A., Teh, Y. W., & Doucet, A. (2022). Coin++: Data agnostic neural compression. arXiv preprint arXiv:2201.12904. [3] Strümpler, Y., Postels, J., Yang, R., Van Gool, L., & Tombari, F. (2022). Implicit neural representations for image compression. European Conference on Computer Vision, 2022. [4] Chen, H., He, B., Wang, H., Ren, Y., Lim, S. N., & Shrivastava, A. (2021). Nerv: Neural representations for videos. Advances in Neural Information Processing Systems, 34, 21557-21568. Skills: C++, Python, Deep Learning framework (Tensorflow or Pytorch), English Keywords: Video, Implicit Neural Representation

Expected outcomes		
Patents	Yes	Proposed solutions could be patented
Publications	Yes	The objective is to publish the finding of this internship
Demos	No	
Software modules	Yes	Software Modules should be developed as part of the research effort

INTERNSHIP OFFER
2024VDO-07

Meta-learning method for dictionary-based INR video compression

LAB	VDO	Group	MetaVideo	PROJECT	SympAI
Main Supervisor	Francois Schnitzer				
Additional supervisors	Pierre Hellier				

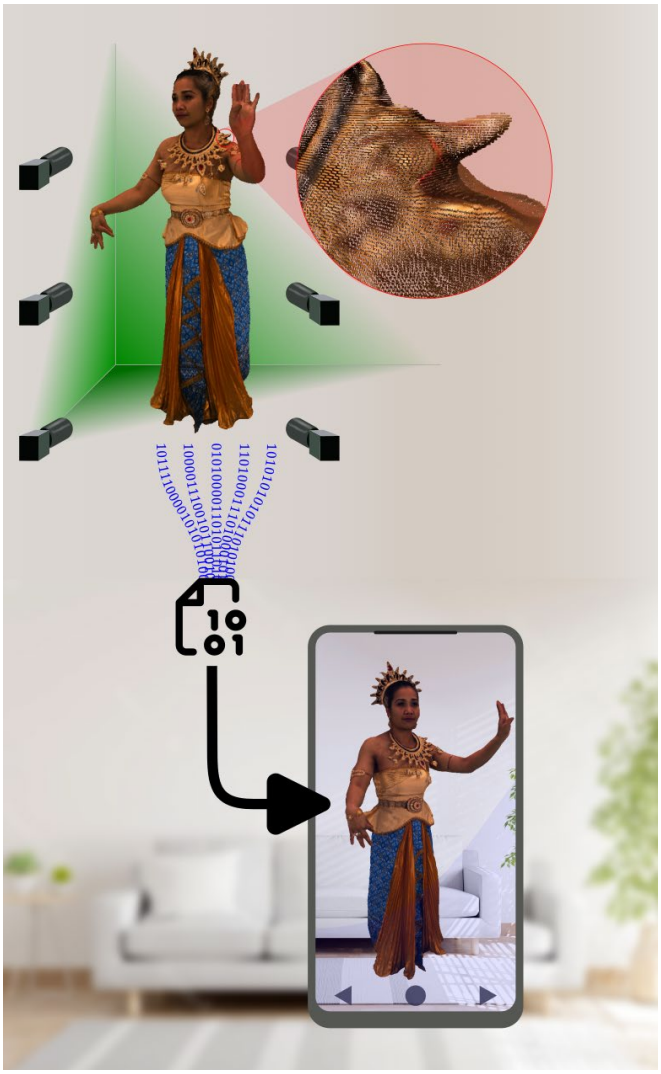
Proposal	
 	<u>Description:</u> Implicit neural representations (INR) have recently emerged as a new tool to model and represent various types of signals, including images [1,2,3] and videos [4]. In particular for image and video compression, these implicit representations can pave the way towards a radically new paradigm: once overfitted on an image to be encoded, transmitting the image amounts to transmitting the weights of the network, and the decoding is a straightforward evaluation of the neural network for each decoded pixel. Although promising, these methods [2,3,4] have not yet reached the performance of current state-of-the-art traditional and learning based approaches. This internship proposes to study further to improve these implicit representations for image and video compression. More specifically at InterDigital, we develop dictionary-based methods to tackle the spatial and temporal redundancy of the signal. The Internship proposes to improve the encoding efficiency using a known dictionary of implicit functions. In particular, meta learning strategies [5] will be investigated to learn the optimal sparse coding technique. A first strategy to develop is to build a dataset of observed patch, alongside with their sparse coded representations, and train a neural network to perform the sparse optimization as a single fast-forward pass of the network. That will eventually lead to an efficient encoding, compatible with broadcast context. References

	[1] Sitzmann, V., Martel, J., Bergman, A., Lindell, D., & Wetzstein, G. (2020). Implicit neural representations with periodic activation functions. Advances in Neural Information Processing Systems, 33, 7462-7473. [2] Dupont, E., Loya, H., Alizadeh, M., Goliński, A., Teh, Y. W., & Doucet, A. (2022). Coin++: Data agnostic neural compression. arXiv preprint arXiv:2201.12904. [3] Strümpler, Y., Postels, J., Yang, R., Van Gool, L., & Tombari, F. (2022). Implicit neural representations for image compression. European Conference on Computer Vision, 2022. [4] Chen, H., He, B., Wang, H., Ren, Y., Lim, S. N., & Shrivastava, A. (2021). Nerv: Neural representations for videos. Advances in Neural Information Processing Systems, 34, 21557-21568. [5] VANSCHOREN, Joaquin. Meta-learning. Automated machine learning: methods, systems, challenges, 2019, p. 35-61 (https://library.oapen.org/bitstream/handle/20.500.12657/23012/1/1007149.pdf). Skills: C++, Python, Deep Learning framework (Tensorflow or Pytorch), English Keywords: Video, Implicit Neural Representation

Expected outcomes		
Patents	Yes	Proposed solutions could be patented
Publications	Yes	The objective is to publish the finding of this internship
Demos	No	
Software modules	Yes	Software Modules should be developed as part of the research effort

INTERNSHIP OFFER
2024VDO-11
Encoding of texture for 3D images/videos represented by point clouds

LAB	VDO	Group	Immersive	PROJECT	G-PCC
Main Supervisor	Gustavo Sandri				
Additional supervisors	Franck Thudor, Bertrand Chupeau				



Description:

3D images/videos are captured from the real world using an array of cameras. The images captured by each camera are combined to create a point cloud representing the outer shell of objects in the scene. This point cloud can be further processed, creating a mesh representation.

Our aim is in the encoding of such data to be stored or transmitted and used for Virtual Reality, Augmented Reality, Telepresence, etc. We work directly with the point cloud aiming real time applications where the time to compute the mesh representation may be prohibitive.

The point cloud is composed of two components: the geometry, that is the (x,y,z) position of the points in space, and the attribute, that describes the texture by associating a (r,g,b) color to each point. In this internship we will focus on how to compress the colors associated to the points, referred to as attribute encoding.

The attribute encoding is done in 3 stages:

1. Prediction of color based on points already encoded, either in the current frame or on previous frames;
2. Transform of the residue between the points' colors and the predicted colors. The transform needs to take advantage of the low pass characteristic of real images;
3. Quantization and entropy encoding of the transformed coefficients.

The objective of this internship will be to test and propose new transforms (stage 2) for the attribute encoding of 3D images/videos represented by point clouds. The result of this internship may be proposed to be adopted by the MPEG standard of point cloud coding.

You will be working in the Geometry-based Point Cloud Compression team, in the Video Lab located in Rennes, FRANCE.

Skills:
Graduate student (M.Sc.) in Computer Engineering, Computer Science, Software Engineering, or related fields.
Knowledge in one or more of the following:
3D geometry, computer vision, computer graphics, C/C++, python.
Fluent English. Excellent written and verbal communication skills. Strong team player.

Keywords:
Point cloud, virtual reality, augmented reality, video coding, image coding.

Expected outcomes		
Patents	Yes	Implementation of advanced features may lead to identify some limitations / novel approaches
Publications	Yes	Standardization contributions
Demos	No	
Software modules	Yes	C++ functions, python scripts

INTERNSHIP OFFER
2024VDO-12
Internship on 3D Dynamic Mesh Compression

LAB	VDO	Group	VIM	PROJECT	D-MESH
Main Supervisor	Maja Krivokuća				
Additional supervisors	Olivier Mocquard				

Proposal	
	Description: 3D representations evolving over time, such as dynamic 3D meshes, have recently become very popular for a variety of applications such as XR. During a 6 months internship in Rennes, France; the student will learn about state-of-the-art Dynamic Mesh Compression techniques and immerse themselves in the activities of the team. The student will work alongside team members to evaluate a specific aspect and identify potential areas for innovation. Skills: • Master's degree with a background (e.g., university courses and/or previous work) in information theory and data compression. • Previous experience in 3D data (especially mesh/point cloud) compression and/or processing. • Coding competencies in C++ Additional coding skills in a prototyping language such as MATLAB would also be appreciated. • Proficiency in English (for reading/writing scientific documentation), French a plus. Keywords: 3D, Mesh, Compression, Immersive, XR.

Expected outcomes		
Patents	Yes	The student will participate in brainstorm and contribute to IP generation.
Publications	Maybe	Depending on progress, a publication in a renowned conference may be envisaged.
Demos	No	The development of demonstrations is not expected.

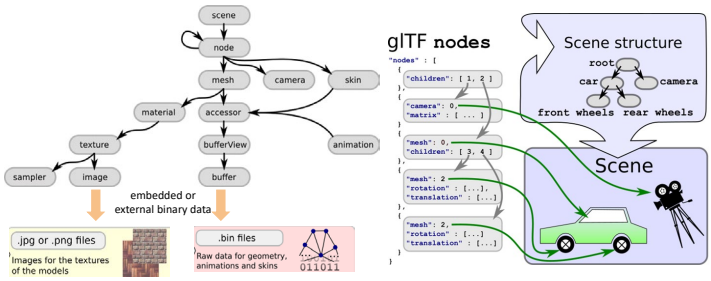
Software modules	Yes	The student will develop code corresponding to their work and integrate it in the software platform.

INTERNSHIP OFFER
2024VDO-13

Elasticity analysis for eXtended Reality (XR) applications over a wireless network

LAB	VDO	Group	Immersive	PROJECT	XR over Wireless
Main Supervisor	Patrice Hirtzlin				
Additional supervisors	Etienne Faivre d'Arcier, Pierrick Jouet, Loic Fontaine, Sylvain Lelievre				

Proposal	
 	Description: Elasticity analysis for eXtended Reality (XR) applications over a wireless network. eXtended Reality (XR) corresponds to technologies that create rich, immersive and interactive virtual worlds. In Augmented Reality (AR) applications, these virtual worlds are seamless inserted into the user's real environment. The framework of an XR experience needs to capture data from the user's real environment through sensors, to update and to render an interactive XR scene composed of numerous and various contents such as 2D or immersive video streams, computer-generated geometries and textures arranged in a scene graph representation. Therefore, a rich XR experience may lead to stringent requirements on data rates and latencies for a wireless network. The objective of this internship will be to analyze what kind of mechanisms and/or profiles can be defined to provide some elasticities on the data rate and/or on the latency constraints for an XR experience. These elasticities can then be used by a XR framework to maintain an acceptable Quality of Experience (QoE) for the user while the wireless network conditions are evolving.



This internship is mainly composed of the three following steps:

- a state-of-the-art analysis on the current mechanisms/profiles impacting the data rate/latency of a XR experience based on a scene graph representation,
- the identification and the definition of the relevant levels of elasticity for typical XR use cases and User Equipment (UE) architectures,
- the implementation of some key mechanisms/profiles into our internal Unity-based XR test bed for validation and demonstration.

You will be working in a team of Mixed Reality experts, in the Video Lab located in Rennes, FRANCE.

Skills:
Graduate student (M.Sc. or Ph.D.) in Computer Engineering, Computer Science, Software Engineering, or related fields.
Knowledge in one or more of the following:
3D geometry, computer vision, computer graphics.
Unity 3D, C#, C/C++, python.
Fluent English. Excellent written and verbal communication skills. Strong team player.

Keywords:
eXtended Reality, scene graph representation, Quality of Experience (QoE), Unity 3D

Expected outcomes		
Patents	Yes	Research-oriented internship with strong potential of patent disclosures as there is a lack of unified mechanisms/profiles to adapt the data rate/latency of XR experiences over a wireless network
Publications	Yes	The XR elasticity concept in the scope of 3GPP QoE is quite

		new
Demos	Yes	Demonstration of some key mechanisms/profiles to manage QoE
Software modules	Yes	Unity plugins, C# scripts

INTERNSHIP OFFER

4K High Dynamic Range Video Processing & Test Tools Development

LAB	VDO	Group	HDR solution	PROJECT	HDR
Main Supervisor	Frédéric Plissonneau				
Additional supervisors	Patrick Morvan / Robin Le Naour /				

Proposal

Description:
Advanced HDR (https://advancedhdrbytechnicolor.com/) is a suite of high dynamic range (HDR) video production, distribution and display solutions that leverages machine learning to maximize the image quality of any HDR format. In the context of our 4K-HDR video studio, we need to enrich our services with a new recording, analysis, and playback tool. It will be used to support our research activities, to guarantee the quality of our deliverables as well as to qualify our partners' equipment.

The objective of the internship is to develop a real-time software “AHDR video player” to drive a professional video interface card (SDI type video link). This software will rely on an SDK (C++) provided by the card manufacturer to implement the playback/recording functions of AHDR videos. In addition to the video, it means extracting the metadata transported in the VBI to store them in a mov-type video container and restore them on the SDI video link in phase with the video during playback.
Advanced features will also be considered such as automated verification by combining image injection and on-the-fly recording of outputs & porting of a real-time HDR video processing solution based on another SDI interface card.

The intern will work in a team of video processing research and development experts. He/She will benefit from their expertise to discover research and technology development, validation and testing before transfer to partners and equipment vendors. During the internship, he/she will be able to participate to demonstrations and will be integrated into the HDR project team.	
Skills : C++, Microsoft Visual, image/video processing is a plus	
Keywords : real-time software , UHD, HDR, Video	

Expected outcomes		
Patents	no	
Publications	no	
Demos	maybe	
Software modules	Yes	Enhancement of the services offered by the 4K-HDR video studio

INTERNSHIP OFFER

4K High Dynamic Range Video implementation with GPU

LAB	VDO	Group	Video Solutions	PROJECT	HDR
Main Supervisor	Nicolas Caramelli				
Additional supervisors	Robin Le Naour				

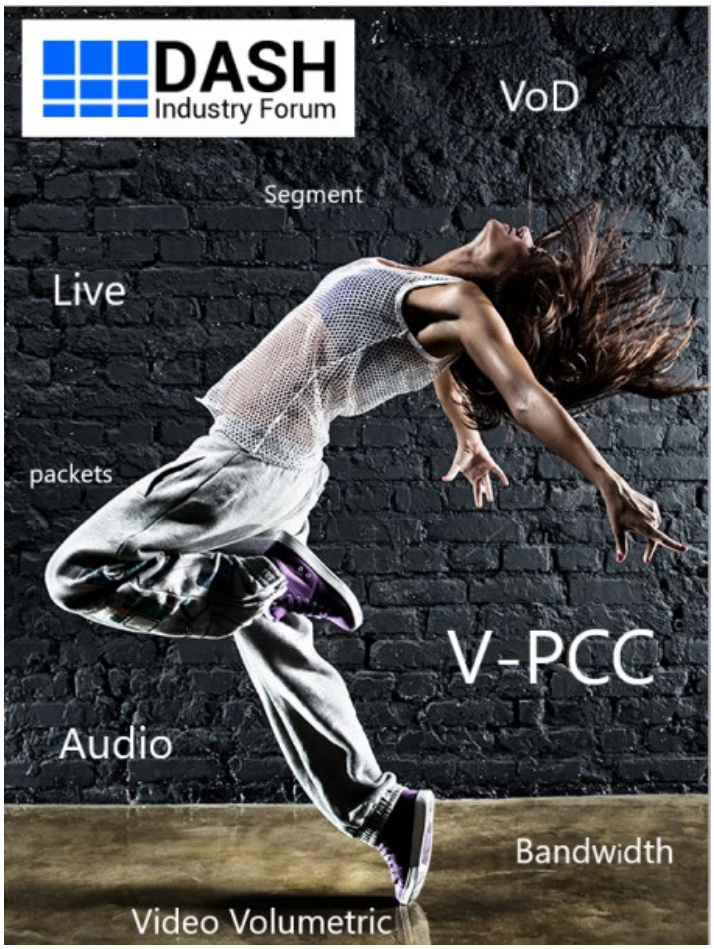
Proposal	
Advanced HDR by technicolor OpenGL ES Vulkan	**Description**: *Context*: High Dynamic Range (HDR) is a new way to capture, produce, distribute, and render images with increased brightness, deeper contrast, and more colors. This technology has been implemented for years by many companies and several standards are in place. At InterDigital we support and implement the Advanced HDR by Technicolor solutions (AHDR) and we provide creation and rendering tools to products manufacturers and service providers. *Objective*: Provide a GPU (Vulkan, OpenGL) implementation of the HDR deliveries developed in the team. The intern will work in a team composed of video and color experts and computer engineering experts. The intern will be part of the Video solutions group addressing HDR, 2D video and immersive video solutions. The intern will contribute to the development of GPU based AHDR encoding tools. To do so, he will be tasked to implement real time video processing algorithms in GPU. Implementation will be integrated in a production workflow and benchmarked. Work will be presented to partners and potentially during trade shows demonstrations. Skills: The intern must be familiar with C++ language and GPU programming. Keywords: HDR, image processing, GPU programming, Vulkan, OpenGL, C++

Expected outcomes		
Patents	No	
Publications	Maybe	GPGPU conference
Demos	Yes	Comparison benchmarks with CUDA / OpenCL / OpenGL
Software modules	Yes	Vulkan support to our internal cross-platform library used for GPGPU HDR development

INTERNSHIP OFFER

Immersive Video Live streaming deployment using DASH

LAB	VDO	Group	Video Solutions	PROJECT	Immersive Video
Main Supervisor	Cyril Quinquis				
Additional supervisors					

Proposal	
	Introduction of the Context: DASH streaming is widely used in the delivery of on-demand video content (VoD mode) and also for live events (live mode). In this context, new video codec standards are emerging to compress more efficiently 2D and Volumetric video such as VVC (Versatile Video Compression) for 2D and V-PCC (Video-based Point Cloud Compression) for Volumetric video. Description of the Internship: This internship consists in developing and/or integrating solutions for Live DASH streaming, from the data encoding, the ingest process, the origin server, the DASH server and finally to the client. To that end, the intern will rely on professional equipment but also existing software solutions developed in our Video lab. He/she will study standards around DASH transport but also existing publications/papers. He/she will study the State of the Art and do reasoned proposals. He/she will propose Software architectures that will be discussed with his/her supervisor. The implemented solutions will be documented and evaluated on our platform for 2D and Immersive video streaming and decoding. Skills: video compression, transport protocol, software development in C/C++, python, scripting, autonomy, English. Keywords: MPEG DASH, VVC, Point Cloud Compression

Expected outcomes		
Patents	no	Comments/description
Publications	no	Comments/description
Demos	yes	Comments/description
Software modules	Yes	Comments/description

INTERNSHIP OFFER

4K High Dynamic Range Video Codec Real time implementation

LAB	VDO	Group	HDR solution	PROJECT	HDR
Main Supervisor	Robin le Naour				
Additional supervisors	Angelo Mazzante				

Proposal	
Advanced HDR by technicolor	Introduction: Context High Dynamic Range (HDR) is a new way to capture, produce, distribute, and render images with increased brightness, deeper contrast, and more colors. This technology has been implemented for years by many companies and several standards are in place. At InterDigital we have standardized and developed the Advanced HDR by Technicolor (AHDR) and we provide creation and rendering tools to products manufacturers and service providers. *Objective: Enhance current real-time encoder developed in the team to support various platforms and architectures.* *Details*: The intern will be part of the HDR team and will have the mission to enhance the video coding library. The first part of the internship will be to convert/implement current SIMD code (SSE/AVX2/AVX512) to an instruction set abstract version using an open-source library. The intern will have to ensure no performance regression will occur and that results will remain consistent with the former code version, to achieve this the intern may have to implement unit test, benchmarks, and a validation toolchain with: -Gitlab CI/CD

	-memory checker -syntax analyzer -functional test -deployment (packaging/deployment) The intern will work in a team of video processing research and development experts. He/She will benefit from their expertise to discover research and technology development, validation and testing before transfer to partners and equipment vendors. During the internship, he/she will be able to participate to demonstrations and will be integrated into the HDR project team. Keywords: SIMD, AVX, neon, HDR devops, CICD, GitLab

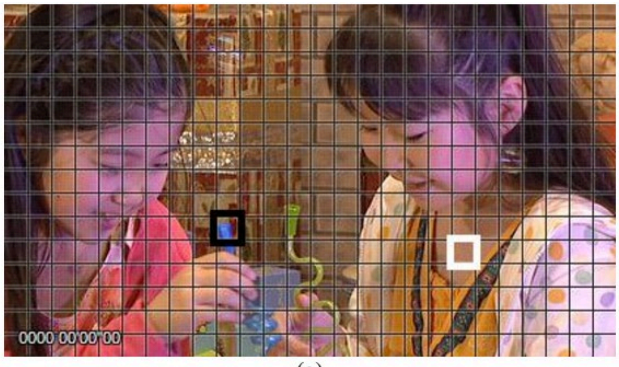
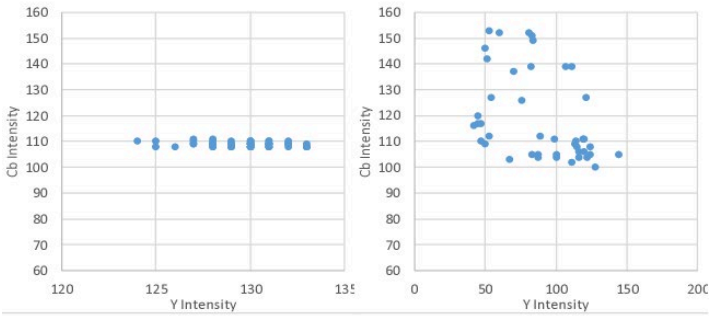
Expected outcomes		Unified SLHDR SDK and apple capability
Patents	no	
Publications	no	
Demos	yes	
Software modules	Yes	SIMD dynamic dispatch, Enhanced production toolchain

INTERNSHIP OFFER

2024VDO-21

Exploration of improved video coding technologies

LAB	VDO	Group	V2S	PROJECT	NGVC
Main Supervisor	Philippe Bordes				
Additional supervisors	Franck Galpin, Fabrice Urban				

Proposal	
 (a)  (b) (c) Fig. 3 Examples of the diversity of cross-component relationship.	The internship will take place in a research team focused on 2D video coding. The team develops coding technologies to improve video compression. The team is deeply involved in video coding standardization (MPEG, JVET). JVET has developed the state-of-the-art standard named VVC, and started exploring enhancement technologies once VVC standard had been issued. The purpose of the internship will be to investigate such enhancements above VVC performance. The study will be based on existing works already achieved in the team about coding of the color components, using cross-component prediction tools. The internship will explore normative as well as encoder-only improvements. **Skills:** - knowledge of C++ - knowledge of video coding standards (e.g. HEVC, VVC) appreciated - interest for algorithm research **Keywords**: video coding, VVC, image processing, video processing

	JVET document repository: https://jvet-experts.org/

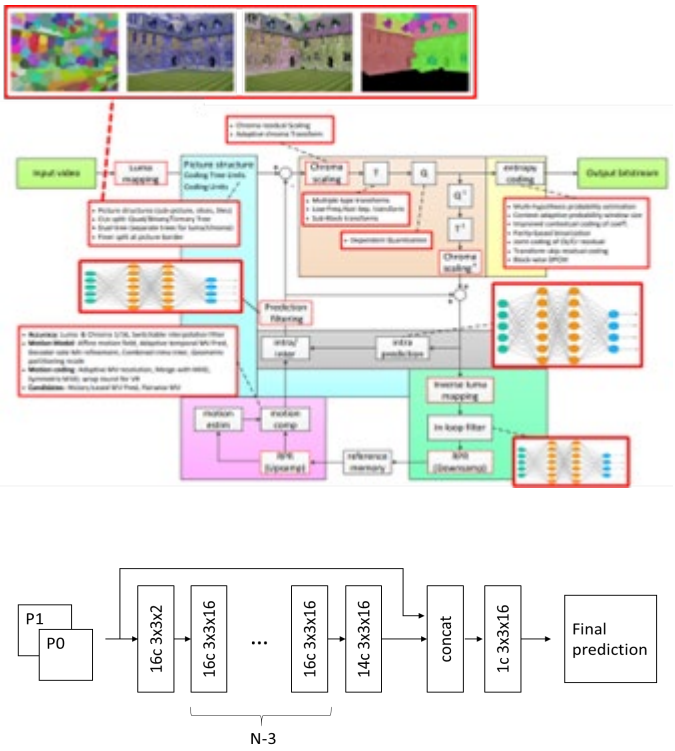
Expected outcomes		
Patents	Maybe	Will depend on the implementation adaptations made during the internship
Publications	Yes	JVET contribution, conference paper possible
Demos	No	
Software modules	Yes	C++ module

INTERNSHIP OFFER
2024VDO-22

Neural Network based inter-prediction blending

LAB	VDO	Group	V2S	PROJECT	NGVC
Main Supervisor	Franck Galpin				
Additional supervisors	Thierry Dumas				

Proposal



InterDigital is looking for a self-motivated intern to work closely with scientists on deep learning for video coding. The prospective candidate is expected to develop a deep learning-based module and integrate it into a reference software developed by the standardization organization JVET, named NNVC. A scientific publication is strongly encouraged and will be a key asset in the candidate's CV. The duration of the internship will be 4-5 months.

Description:
The main goal of the internship is to integrate an NN-based inter-blending process into the NNVC software, and to adapt and improve its design.
NN-based inter-blending was proposed in an early version (see JVET-V0076). It is proposed in this internship to integrate it back in the new software NNVC (see JVET-AD2019) and to evaluate its performance, following the test process defined by JVET (training, dataset, test protocol etc.).

Skills: C++, Python, TF or PyTorch, video coding
Keywords: video coding, Deep Learning

JVET document repository:
https://jvet-experts.org/

Expected outcomes		
Patents	Maybe	Will depend on the implementation adaptations made during the intership
Publications	Yes	JVET contribution planned, conference paper possible
Demos	No	
Software modules	Yes	Training scripts, C++ module

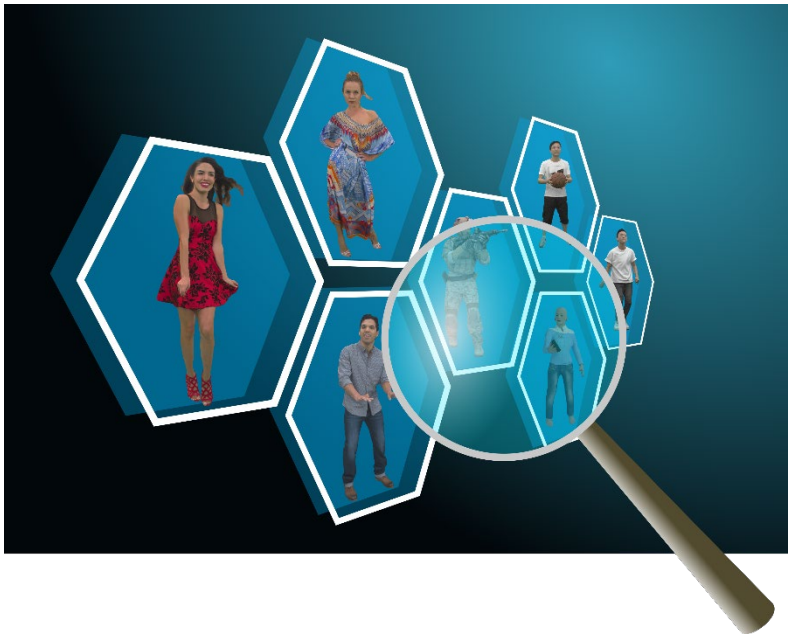
INTERNSHIP OFFER

Development of a framework for subjective testing

LAB	VDO	Group	Immersive	PROJECT	G-PCC
Main Supervisor	Franck Thudor				
Additional supervisors	Gustavo Sandri, Bertrand Chupeau				

Description:

3D images/videos are captured from the real world using an array of cameras. The images captured by each camera are combined to create a point cloud representing the outer shell of objects in the scene. This point cloud can be further processed, creating a mesh representation.



When a point cloud is used for applications like Virtual Reality, Augmented Reality, Telepresence, etc... an encoding step is mandatory to reduce the bulk of the transmitted data. As a result, the rendered point cloud often suffers from artifacts that might lead to a distracting or even uncomfortable experience from a user point of view.

A lot of research has been carried out to assess the quality for the 2D video compression, notably how to correlate objective metrics (simple and cheap) and subjective evaluation (expensive). When it comes to encoding point clouds, only few investigations have been performed. Firstly, because this topic is less mature, and secondly because the way an assessment must be done is not obvious. We propose to study this problem during the internship.

Main steps that are expected during this internship:

1. A deep state of art on the objective/subjective metrics related to point-cloud compression;
2. Analysis of the limitations and constraints for quality assessment;
3. Propose a methodology for the quality assessment;

4. Develop and setup of a complete automatic framework to generate test data, including encoding, decoding, rendering, video generation;
5. Conduct measurement campaigns to validate the framework.

You will be working in the Geometry-based Point Cloud Compression team, in the Video Lab located in Rennes, FRANCE.

Skills:
Graduate student (M.Sc.) in Computer Engineering, Computer Science, Software Engineering, or related fields.
Knowledge in one or more of the following:
3D geometry, computer vision, computer graphics, machine learning, C/C++, python.
Fluent English. Excellent written and verbal communication skills. Strong team player.

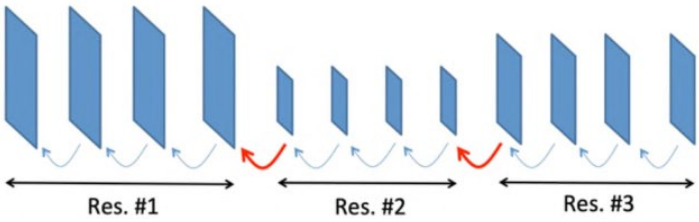
Keywords:
Point cloud, virtual reality, augmented reality, view synthesis, quality assessment.

Expected outcomes		
Patents	Yes	Although not a direct objective of the internship, some novel ideas may arise during the internship period.
Publications	Yes	Standardization contributions, scientific publications
Demos	No	
Software modules	Yes	Python scripts, C++ functions

INTERNSHIP OFFER
2024VDO-24

ML-based Adaptive resolution video coding

LAB	VDO	Group	V2S	PROJECT	NGVC
Main Supervisor	Hassane Guermoud				
Additional supervisors	Kevin Reuzé, Philippe Bordes				

Proposal	
 	The internship will take place in a research team focused on 2D video coding. The team develops coding technologies to improve video compression. The team is deeply involved in video coding standardization (MPEG, JVET). JVET has developed the state-of-the-art standard named VVC, that comprises a coding tool (named Reference Picture Resampling) enabling to adaptively change the resolution of the coded picture. This feature turns out to have high potential coding gains, but needs fine tuning at the encoder side. The purpose of the internship will be to investigate such tuning, based on initial works made by the team using machine learning approaches. **Skills:** - knowledge on ML library (e.g. scikit) or a deep learning framework (Pytorch/Tensorflow) - knowledge of C++ - knowledge of video coding standards appreciated - interest for algorithm research **Keywords**: video coding, VVC, image processing, video processing JVET document repository:

	https://jvet-experts.org/

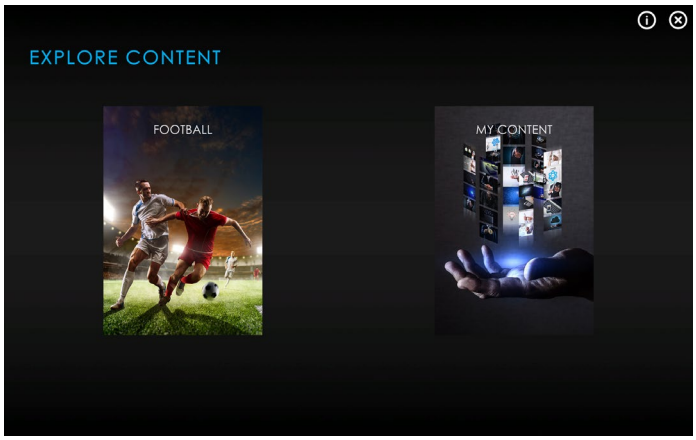
Expected outcomes		
Patents	Maybe	Will depend on the implementation adaptations made during the internship
Publications	Yes	JVET contribution, conference paper possible
Demos	No	
Software modules	Yes	C++ module

INTERNSHIP OFFER

iOS port of Immersive Video Demo with Unreal Engine

LAB	VDO	Group	Video Solutions	PROJECT	Immersive Video
Main Supervisor	Bertrand LEROY				
Additional supervisors	Julien MULARD				

Proposal



Introduction of the Context:
The Video Solutions team has developed a rendering demonstrator on Unity for 2D (VVC & HEVC) and V3C video codecs (MIV & V-PCC) based on a C++ plugin for decoding that can be compiled on Windows or Android. The application illustrating volumetric video content consumption for telelearning use case has been developed with Unity 3D.

Description of the Internship:
This internship consists in porting the plugin to the iOS architecture and the demonstrator to Unreal Engine to demonstrate it on Apple devices (iPhone or Apple TV).
The intern will have to set up different scenes using the C++ volumetric video decoder plugin, with interfaces adaptation if needed for iOS, and illustrate different use cases. The application should run on Windows and Android with AR and VR capabilities, and multiple devices like smartphones and head mounted display (stereoscopic views) thanks to Unreal Engine.
She/he will be fully integrated in the video lab will develop new modules and will produce appropriate documents.
Knowledge in shader development and other game engines (like Godot) will be an added value.

Skills: Unreal Engine, Unity, C# and C/C++ on iOS, shaders, scripting, autonomy, English.

Keywords: iOS, 3D, C++, Unreal.

Expected outcomes		
Patents	No	
Publications	No	
Demos	Yes	The intern may be asked to present the demo at trade shows.
Software modules	Yes	The intern will adapt the existing modules to iOS architecture and port the demo to Unreal Engine.

INTERNSHIP OFFER

Live game capture in engine for volumetric video streaming

LAB	VDO	Group	Video Solutions	PROJECT	Immersive Video
Main Supervisor	Julien MULARD				
Additional supervisors					

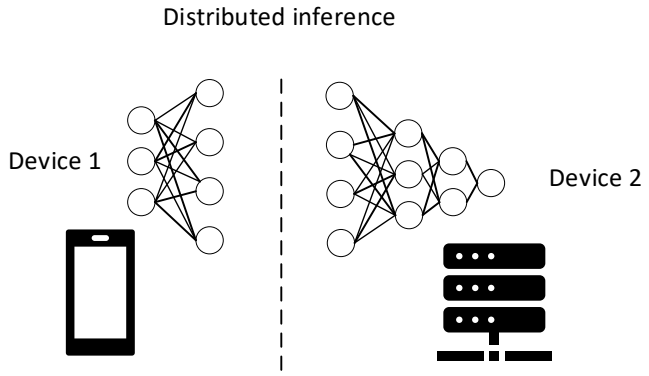
Proposal	
	A volumetric video is comprised of a sequence of frames, and each frame is a static 3D representation of a real-world object or scene capture at a different point in time. The InterDigital Video Solution team has developed a demonstration software capable of decoding both 2D videos (HEVC, VVC) and volumetric videos (V-PCC for point clouds data) in real time on Windows and Android devices. This demonstrator is based on a custom C++ plugin to manage the media decoding, and on Unity 3D for the rendering and applicative side of things. Description of the Internship: This internship will consist in creating a virtual capture rig in engine (Unity 3D, Unreal, Godot) to generate point cloud data in real time from a gameplay sequence. The output of the virtual capture rig will feed a Video-based Point Cloud Compression (V-PCC) encoder to enable streaming of volumetric game sequences. The internship will then work on interfacing the captured volumetric data with an existing real time V-PCC encoder and DASH streaming server. He/She will be working with a team of 2D and volumetric video experts, and will be part of the project team to build an end-to-end demonstration platform. Skills: software development in C/C++, Unity or Unreal or Godot, rendering, shader code, point cloud, bash scripting, autonomy, English. Keywords: Video-based Point Cloud Compression, Unity, Unreal, Godot, Computer Graphic.

Expected outcomes		
Patents	yes	Potential on capture and on interface with VPCC live encoder
Publications	Yes/no	Comments/description
Demos	yes	Part of the V3C end to end demosntrations
Software modules	Yes	Plug-in

INTERNSHIP OFFER

AI for media proof of concept

LAB	Video	Group	Multimedia System	PROJECT	AI for Media
Main Supervisor	Stéphane Onno				
Additional supervisors	Thierry Filoche; Cyril Quinquis				

Proposal	
	Description: This internship involves developing a software demonstrator for an Artificial Intelligence/Machine Learning application (AI/ML) (e.g. Extended Reality) distributed between two devices, particularly focussing on the development of a transport protocol This activity is related to our current AI for Media work within the 3GPP SA4 working group where 3GPP being the standard body for mobile communication, e.g., 5G/6G. Skills: • Master's or PhD with a background in AI or network transport protocols and multimedia systems. • Coding competencies in python, good experience with development tools and scripting • Experience with AI framework, and/or streaming protocols (e.g., MPEG-DASH, RTP/RTCP), including hands-on experience with relevant software libraries and frameworks . • Proficiency in English (for reading/writing scientific documentation), French a plus. Keywords: AI/ML, python, PyTorch, Onnx format, RTP, transport protocols

Expected outcomes		
Patents	yes	The student will participate in brainstorm and contribute to IP generation.
Publications	Maybe	Depending on progress, a publication in a renowned conference may be envisaged.
Demos	yes	The student will integrate their work into the existing demo platform
Software modules	Yes	Same above