Learning-based Point Cloud Compression in JPEG

Stuart Perry (University of Technology Sydney, Australia)

ICME workshop on Immersive Media Compression

Brisbane, Australia, July 10th, 2023

Overview

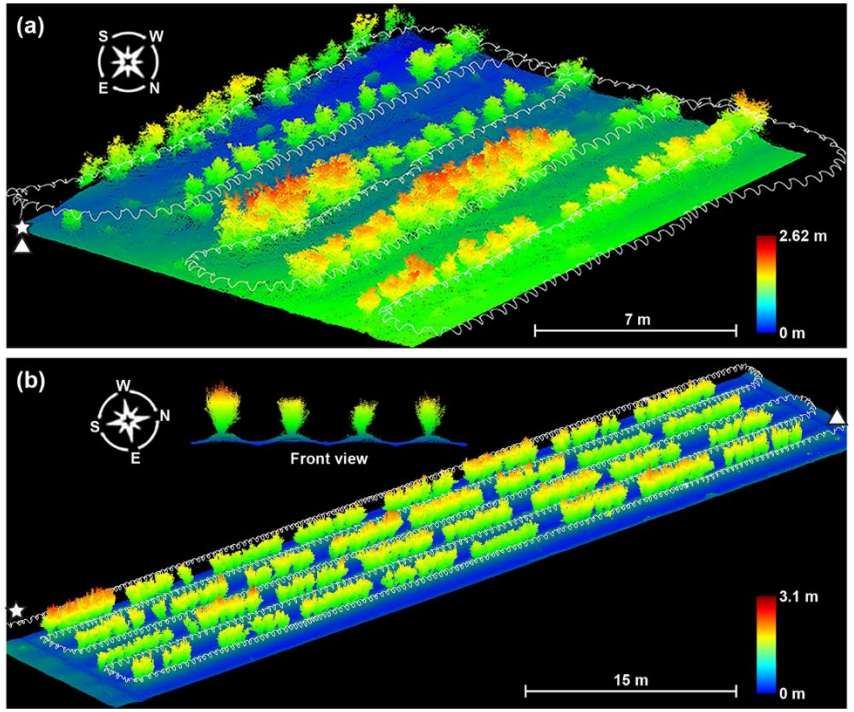
- Overview of JPEG Pleno
- Scope and Stages of JPEG Pleno Point Cloud Activity
- Common Training and Test Conditions
- Results of Final Call for Proposals on JPEG Pleno Point Cloud Coding
- JPEG's current Verification Model
- Next Steps
- Activity Timeline

Rapid Rise of Point Cloud Data

- Point Cloud acquisition and processing is becoming crucial to society
- Consumption of point clouds by machines and humans is growing exponentially
- Standards that support human and machine use are crucial.

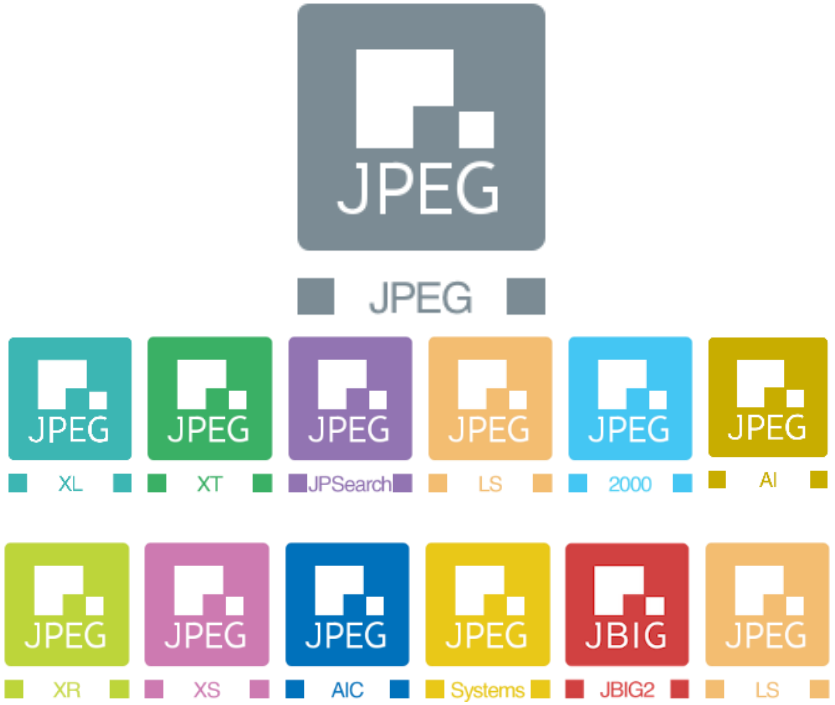
“ByteBridge Launches World’s First Mobile 3D Point Cloud Data Labeling Service”, June 2021
https://medium.com/nerd-for-tech/bytebridge-launches-worlds-first-mobile-3d-point-cloud-data-labeling-service-cf87a4ed2067

“Unlimited 3D Point Cloud Search Cited as Game Changing”, May 2020
https://lidarnews.com/articles/unlimited-3d-point-cloud-search-cited-as-game-changing/



Jiang, Y., Li, C., Takeda, F. *et al.* 3D point cloud data to quantitatively characterize size and shape of shrub crops. *Hortic Res* **6,** 43 (2019). https://doi.org/10.1038/s41438-019-0123-9

What is JPEG Pleno?



JPEG Pleno

- Ad hoc Group on JPEG Pleno - Lightfield
- Ad hoc Group on JPEG Pleno - Holography
- **Ad hoc Group on JPEG Pleno - Point Clouds**

JPEG Pleno is the expansion of the JPEG ecosystem into emerging plenoptic imaging modalities.

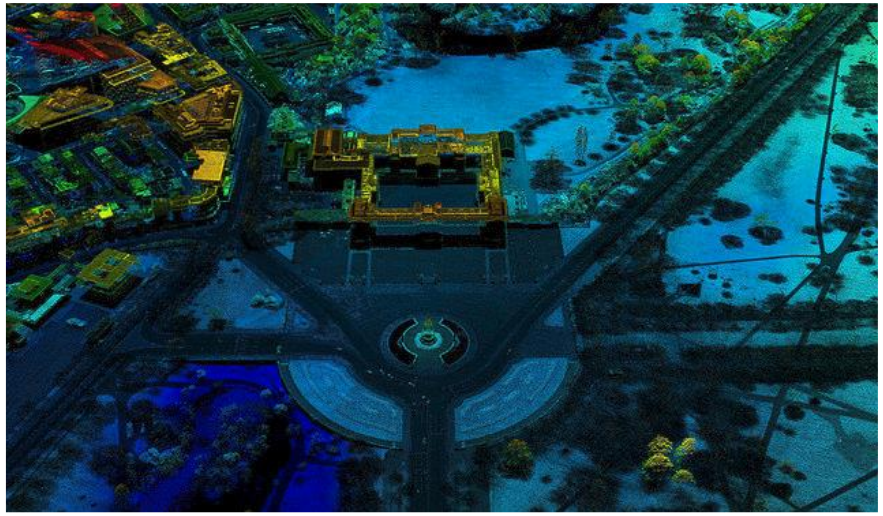
- Light field
- Holography
- Point Clouds

JPEG Pleno is envisioned to provide a holistic approach to the representation of 3D data across these representations.

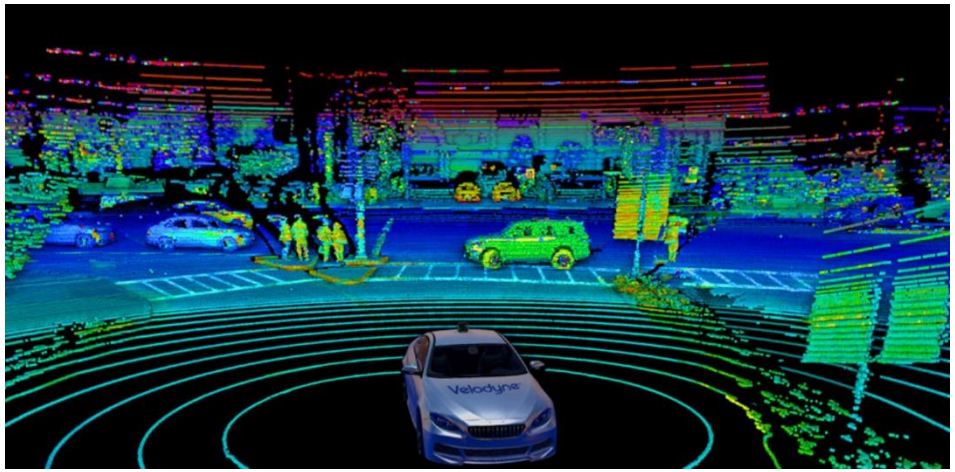
Use Cases and Requirements

Key Use Cases:

- **Wide-area survey/3D mapping**
- **Autonomous driving**
- **Manufacturing – traditional and additive systems**
- **On-line shopping**
- **Fault and defect detection in manufacturing and construction**
- **Cultural Heritage**



The image is a LIDAR scan of Buckingham Palace, UK and is courtesy of Environmental Agency (https://www.flickr.com/photos/environment-agency/27489358013) [CC BY 2.0 (https://creativecommons.org/licenses/by/2.0/)]

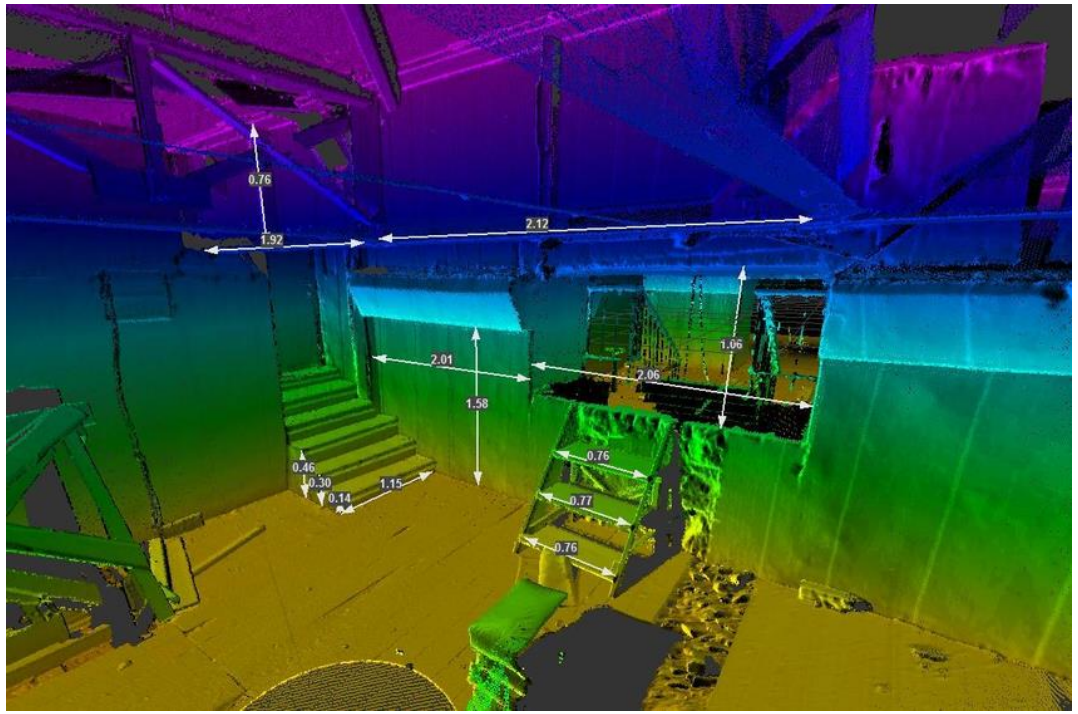


https://www.i-micronews.com/how-lidar-is-getting-ready-for-the-automotive-mass-market-an-interview-with-velodyne/ April 2019

Use Cases and Requirements

Key Requirements:

- **3D Metrology – Preservation of relative point positions**
- **3D Processing:**
  - **Visual Enhancement**
  - **Super Resolution**
- **Computer Vision:**
  - **Object Detection**
  - **Object Classification**
- **Scalability of Geometry and Attributes**
  - **Different degrees of precision, resolution and range**
- **Random Access – Selective decoding of a portion of the point cloud independently of the rest**



Chapter "Positioning and Applications" in National Report for the IAG of the IUGG 2011–2014, June 2015, 10.2205/2015IUGG-RU-IAG

Scope of the Activity

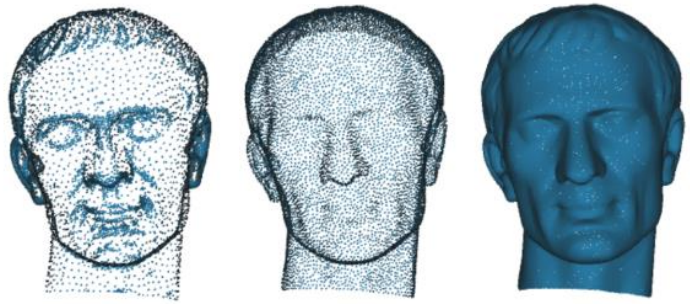
The scope of the JPEG Pleno Point Cloud activity is the creation of a learning-based coding standard for point clouds and associated attributes, offering **a single-stream, compact compressed domain representation, supporting advanced flexible data access functionalities**. This standard targets **both interactive human visualization**, with competitive compression efficiency compared to state of the art point cloud coding solutions in common use, **and effective performance for 3D processing and machine-related computer vision tasks**, and has the goal of supporting a **royalty-free baseline**.

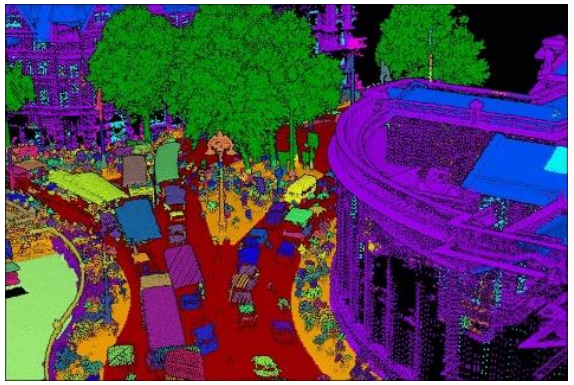
Stages of the Activity

Stage 1: A learning-based coding standard addressing human visualization and **decompressed/reconstructed domain 3D processing and computer vision tasks**;

Stage 2: A learning-based coding standard additionally supporting **compressed domain 3D processing** such as visual enhancement and super-resolution and;

Stage 3: A learning-based coding standard additionally supporting **compressed domain computer vision** tasks such as classification, recognition and segmentation.





Common Training and Test Conditions

Common *Training and* Test Conditions must adapt to a learning-based paradigm to support CfP.

This includes:

- Datasets split into training and testing
- Subjective testing procedures and objective measures must adapt to a rapidly evolving field.
- New objective measures such as computational efficiency need to be considered.

Test Set

- 20 point clouds
  - Chosen based on content diversity
  - Currently under revision!

























Final Call for Proposals on JPEG Pleno Point Cloud Coding

Released at end of 94th JPEG Meeting, 21st January 2022

This call addresses learning-based coding technologies for static point cloud content and associated attributes with emphasis on both human visualization and decompressed/reconstructed domain 3D processing and computer vision with competitive compression efficiency compared to point cloud coding standards in common use, with the goal of supporting a royalty-free baseline.

Results of Call for Proposals

Three submissions to the Call for Proposals:
3 geometry + colour codecs, 2 geometry only codecs

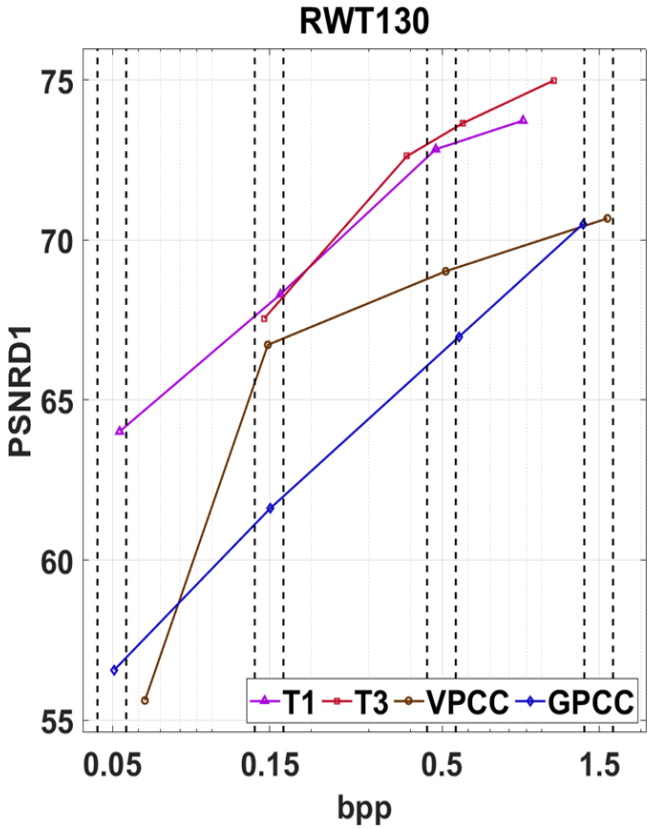
- IT-IST-IPLeiria, Portugal
- EPFL, Switzerland
- University of Science and Technology China, China

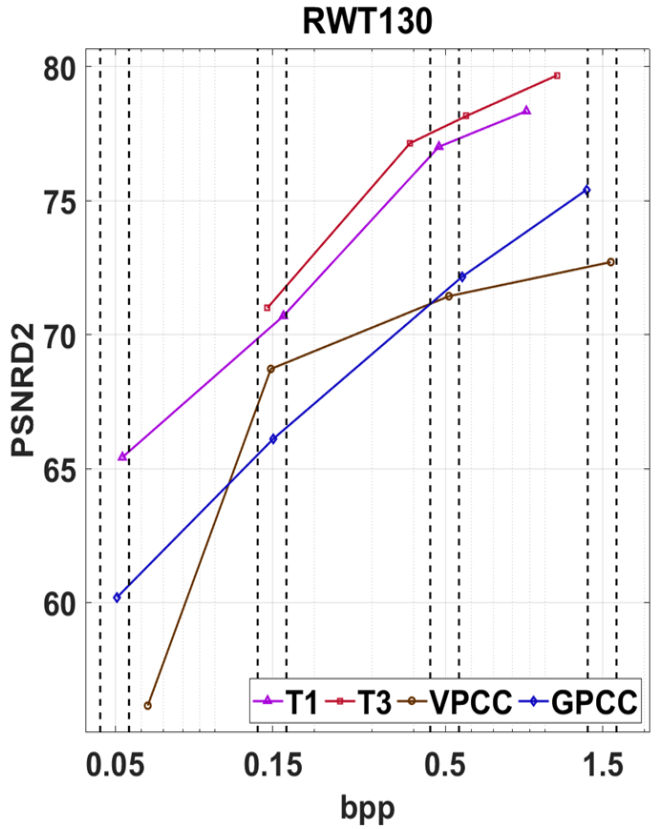
Proposals reviewed at the 96th JPEG Meeting (25-29 July 2022) based on subjective quality assessment and objective metrics as well as computational complexity information

Results of Call for Proposals

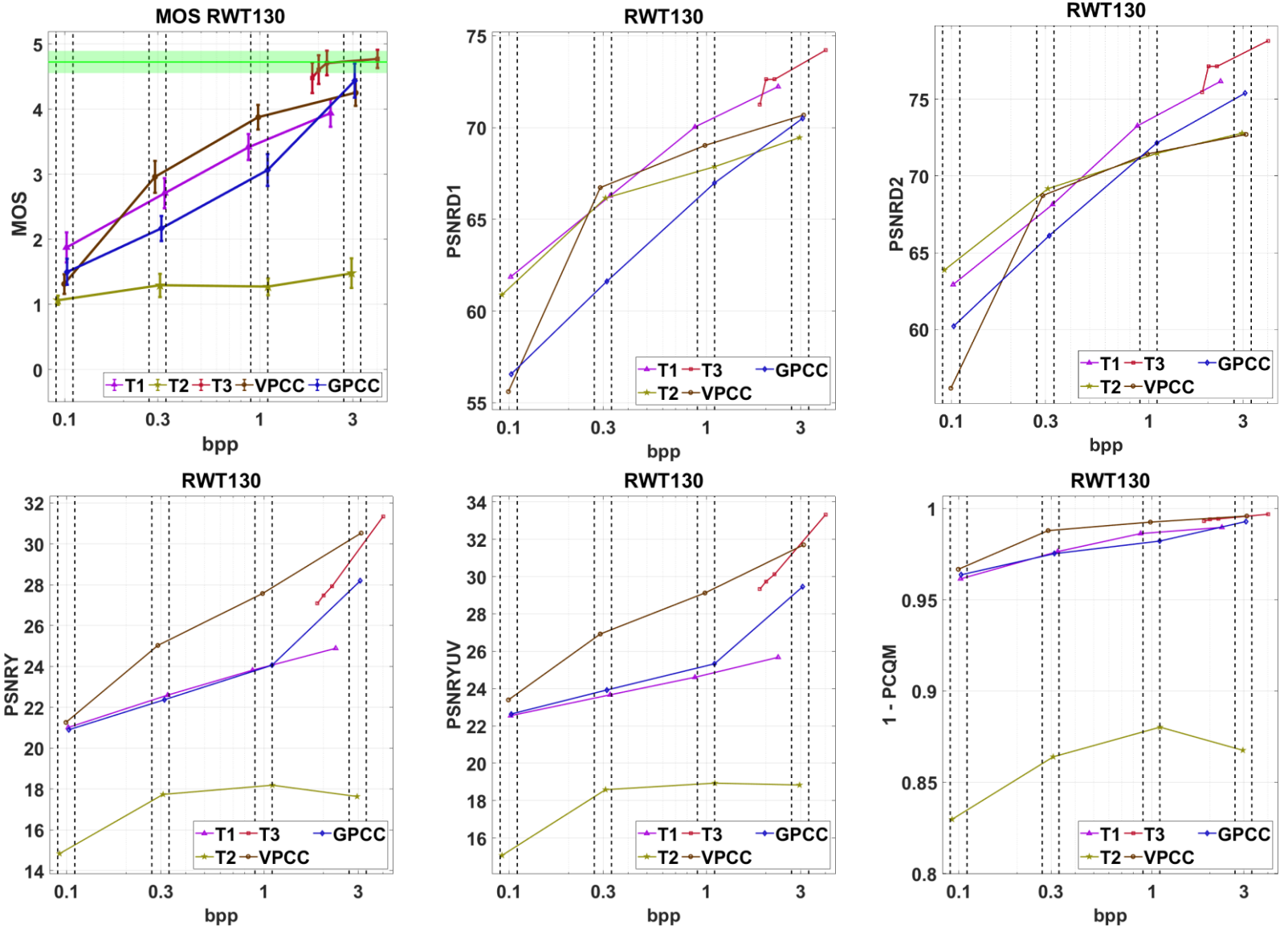
Selection of a proposal (Team 1) as the basis of a Verification Model for the activity

- Strong subjective assessment results compared to other eligible proposals
- Geometry encoding for T1 is often considerably better than anchors.

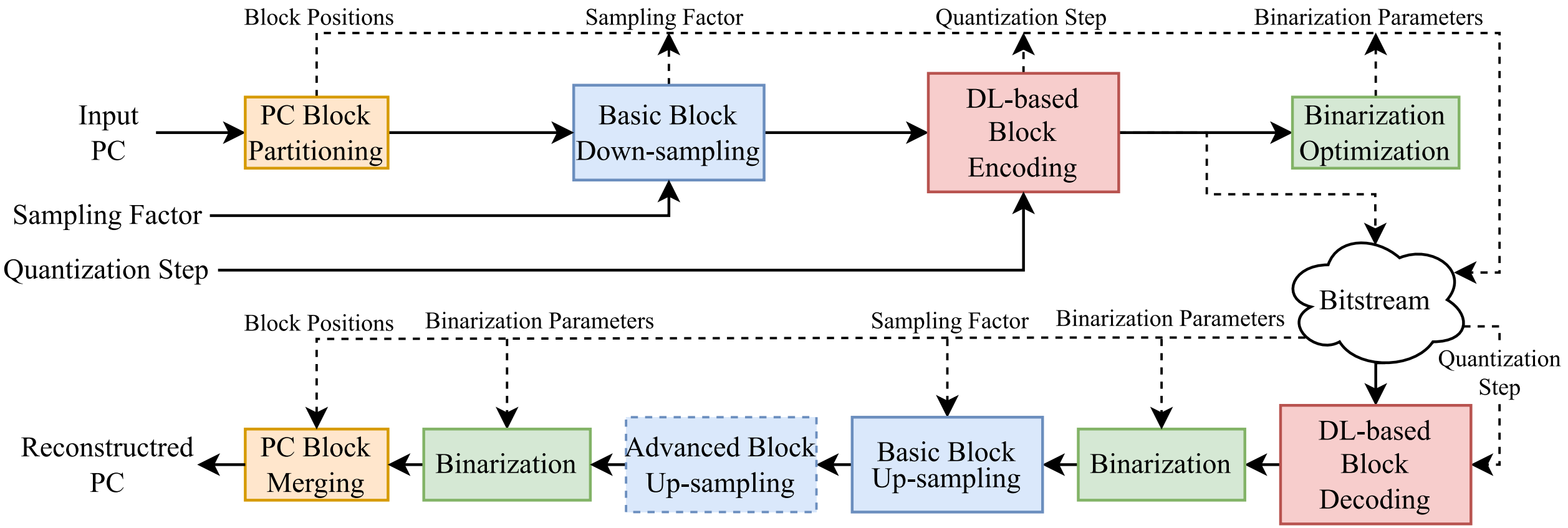




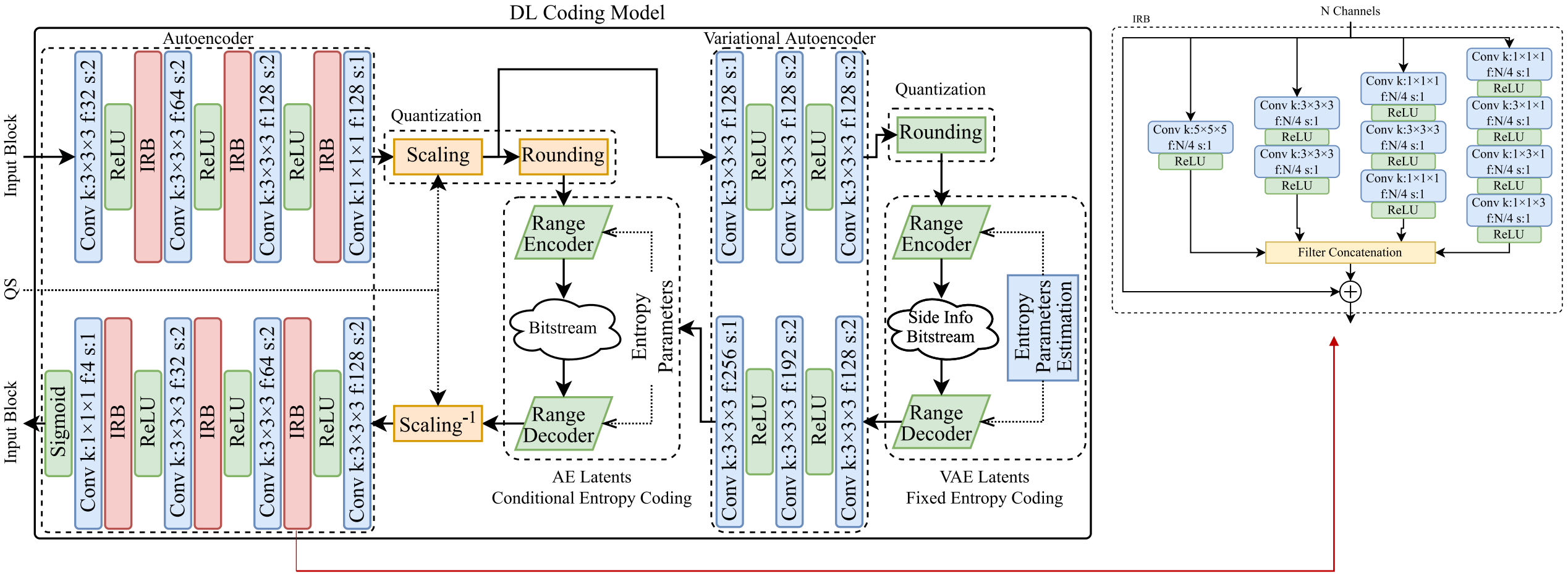
RWT 130 Point Cloud MOS Results (Colour + Geometry)



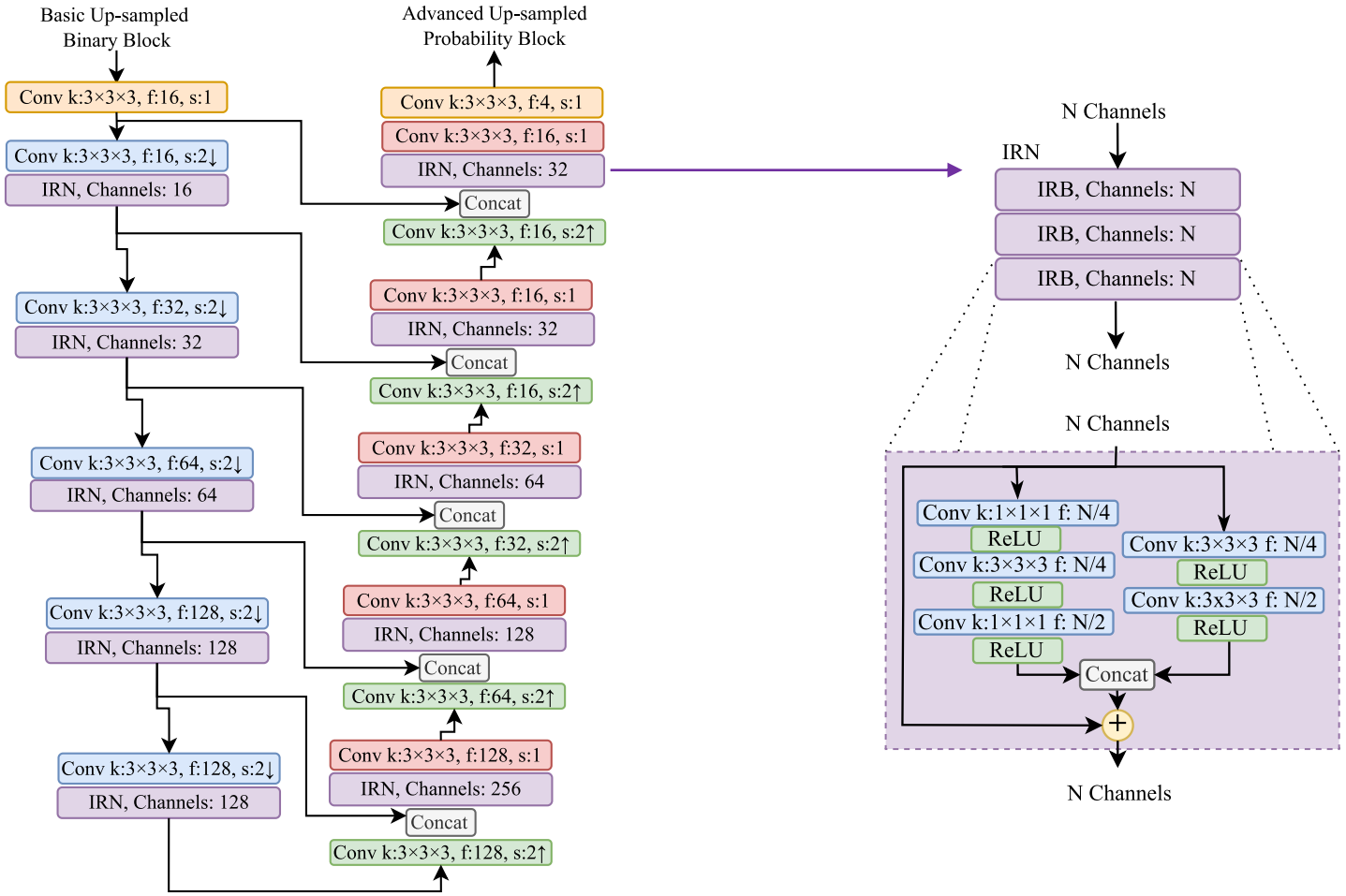
Proposal Selected as Initial Verification Model under Consideration



VM v1: DL Coding Model



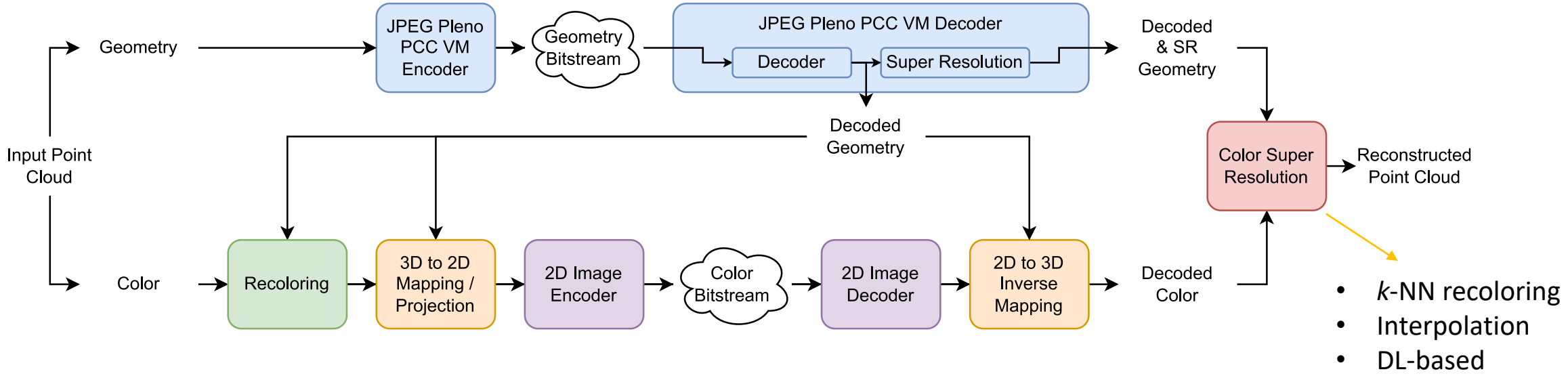
VM v1: DL Advanced Block Up-sampling Model



Current Status of VM

- We have now entered the collaborative phase of the activity:
  - Request for subdivision of ISO/IEC 21794 into Part 6 - JPEG Pleno: Learning-based Point Cloud Coding
  - Validation of VMuC converted from Tensor Flow to Pytorch
  - **Improvements to the colour coding performance through the separation of geometry and colour coding pipelines**

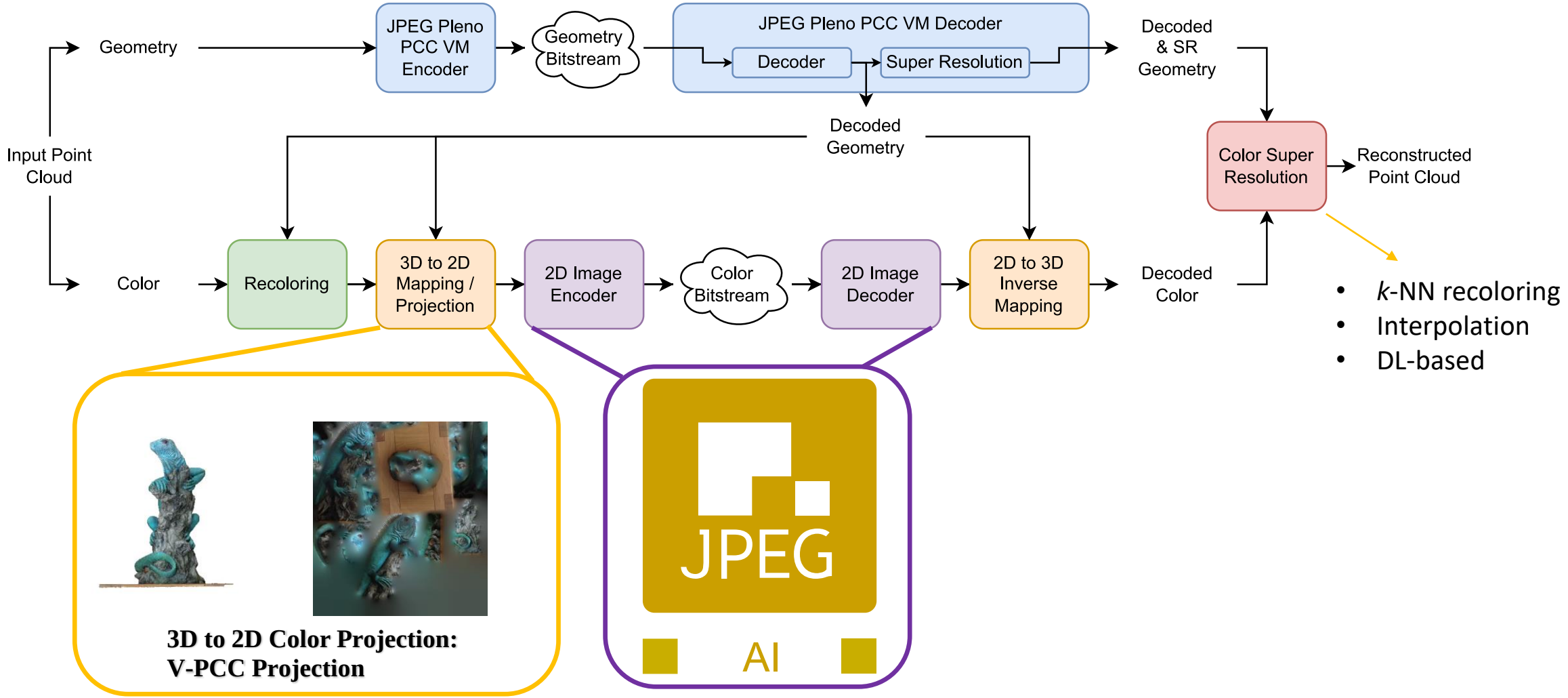
Current VM Architecture



⋆ Approach

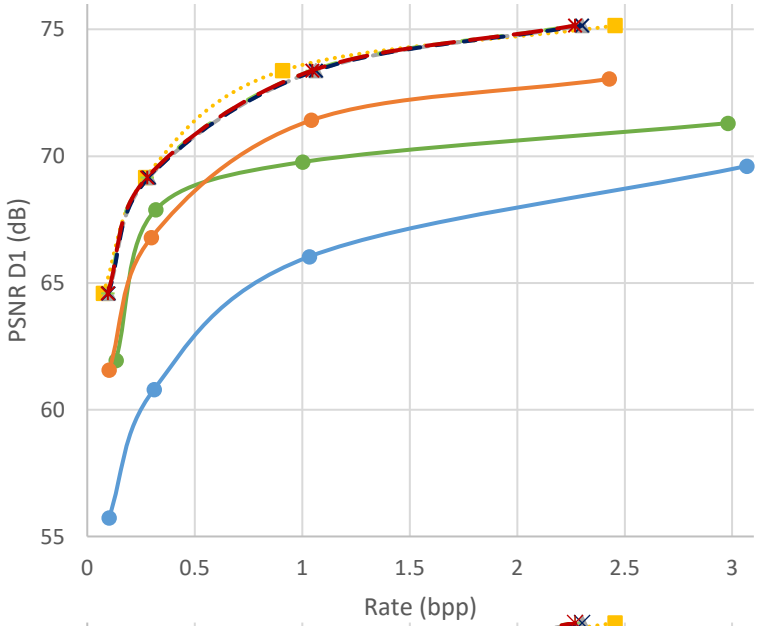
- PC geometry is coded with JPEG Pleno PCC VM
- After recoloring the decoded geometry, the PC color is mapped or projected onto a 2D image
- The 2D color image is coded with a 2D codec
- Super-resolution (SR) is performed separately

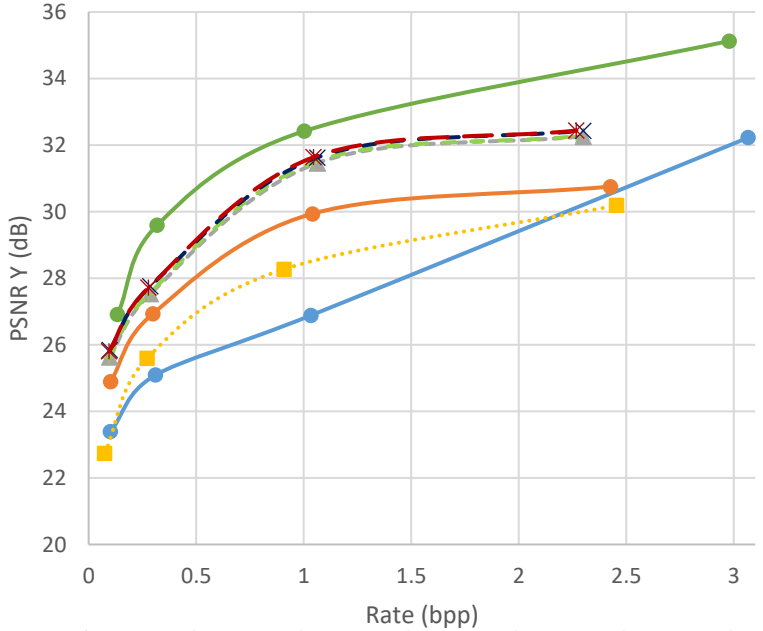
Current VM Architecture

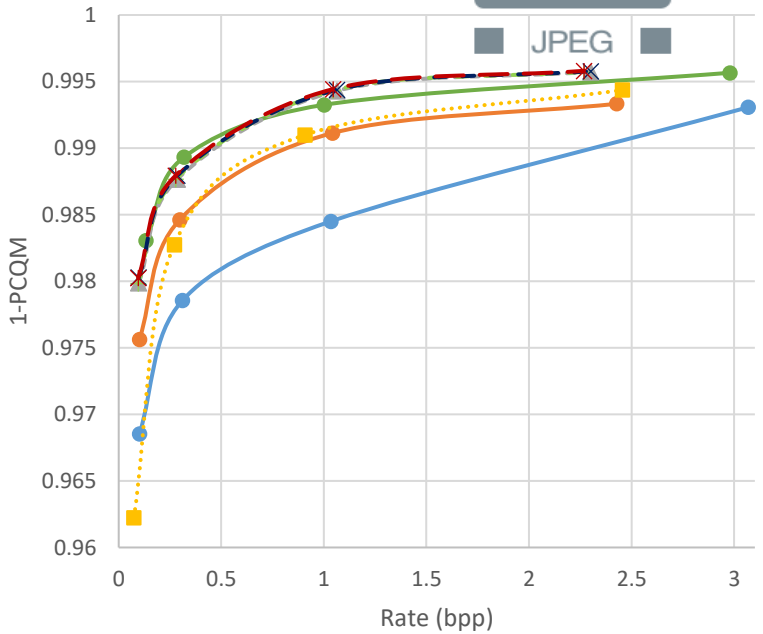


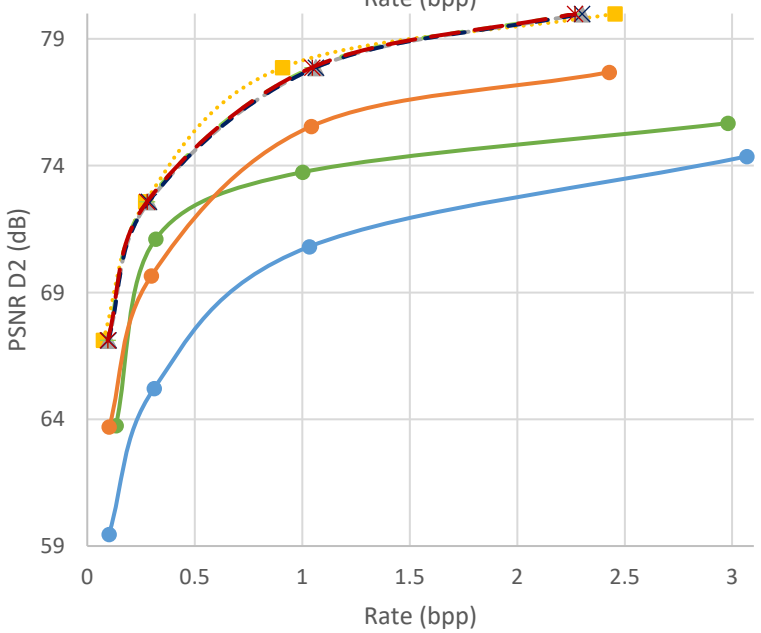
Guarda *et al.*, wg1m99099 – IT/IST/IPLeiria Report on JPEG PCC Core Experiment 3.3

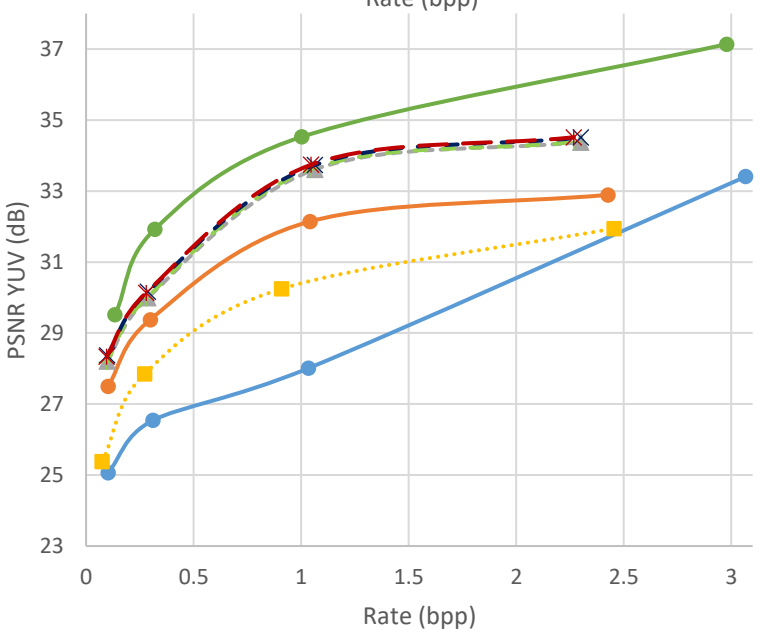
Average RD Performance with Recoloring and SR

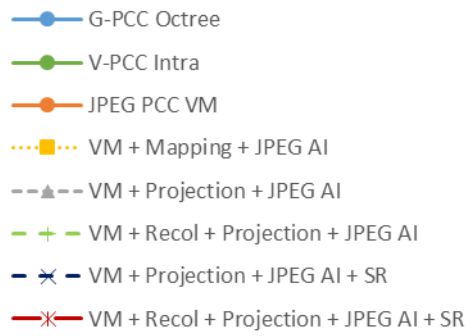










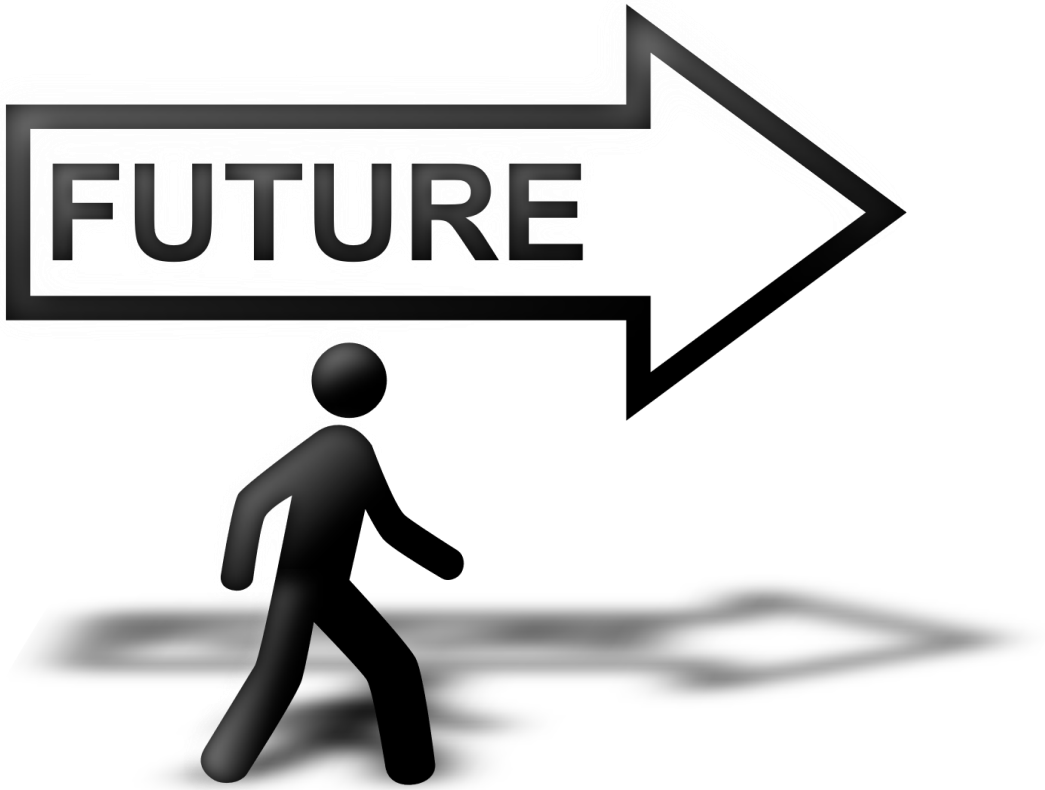


RD Performance: VM + Proj + JPEG AI + Recoloring + SR

	Reference: JPEG Pleno PCC VM									
	PSNR-D1		PSNR-D2		PSNR-Y		PSNR-YUV		PCQM	
PC Name	BD-Rate	BD-PSNR	BD-Rate	BD-PSNR	BD-Rate	BD-PSNR	BD-Rate	BD-PSNR	BD-Rate	BD-PSNR
DinoSkull	-48.3%	2.43	-47.6%	2.86	-8.8%	0.41	-6.2%	0.31	-31.5%	0.002
Lakoon	-51.8%	2.73	-51.4%	3.24	-41.2%	2.04	-36.0%	1.57	-47.0%	0.003
Van	-49.5%	2.42	-48.4%	2.83	-25.7%	0.98	-15.6%	0.57	-37.7%	0.003
StMichael	-46.5%	2.15	-43.2%	2.30	-8.6%	0.31	-7.0%	0.25	-26.9%	0.002
CowStatue	-36.5%	1.74	-33.2%	1.92	-33.1%	0.87	-37.9%	0.92	-50.4%	0.003
Bouquet	-55.0%	3.05	-57.7%	3.74	-34.8%	1.15	-43.1%	1.36	-48.9%	0.005
CatStatue	-68.3%	2.91	-63.4%	3.30	-42.2%	1.38	-40.7%	1.22	-59.8%	0.004
BodyScanBlueShirt	-55.9%	3.14	-54.4%	3.67	-22.3%	0.78	-48.6%	1.39	-52.0%	0.003
Cabbage	-75.3%	4.30	-72.6%	5.09	-57.0%	3.53	-59.6%	3.18	-62.5%	0.006
SteamEngine	14.2%	-0.08	11.9%	0.03	-27.4%	0.75	-31.9%	0.81	-54.1%	0.003
ArmChair	-60.8%	2.13	-62.2%	2.69	-51.8%	2.06	-48.4%	1.75	-64.0%	0.004
BoatJosefa	-31.4%	1.84	-32.5%	2.12	-40.5%	1.11	-41.0%	1.13	-44.8%	0.003
CapitolineWolf	-49.5%	2.38	-48.8%	2.88	-51.9%	2.62	-45.0%	2.03	-56.7%	0.003
KingCrab	-40.8%	2.21	-40.5%	2.54	-16.1%	0.50	-13.4%	0.37	-37.3%	0.003
WoodenChest	-62.7%	3.75	-63.7%	5.04	-82.7%	4.94	-82.9%	4.22	-83.2%	0.006
BodyScanOlia	-56.3%	3.50	-55.4%	4.10	12.0%	0.81	1.9%	0.85	-56.2%	0.003
PaintedEgg	-57.1%	3.62	-54.7%	3.79	-59.0%	2.53	-59.5%	2.37	-76.1%	0.007
Annibal	-53.9%	2.83	-55.0%	3.48	-44.4%	1.81	-43.2%	1.51	-50.7%	0.003
Iguana	-44.8%	3.09	-44.0%	3.31	-26.8%	0.88	-25.7%	0.77	-49.8%	0.003
Pliers	-49.3%	2.93	-51.6%	3.77	-75.5%	3.58	-75.5%	3.25	-74.4%	0.006
goat_skull	-38.4%	1.23	-41.5%	1.94	-37.8%	0.64	-52.1%	0.97	-45.4%	0.003
kinfudesk	-31.6%	1.51	-15.5%	0.95	-16.2%	0.15	-18.4%	0.26	5.5%	-0.001
kinfubooks	-54.1%	2.75	-52.8%	3.79	-25.8%	0.47	-29.7%	0.50	-30.6%	0.002
LivingRoom	-33.3%	0.85	-40.7%	1.32	31.8%	-0.35	25.5%	-0.23	20.6%	-0.001
RuaDeCoimbra	-61.4%	2.96	-63.9%	3.94	-30.3%	0.72	-45.8%	1.21	-56.0%	0.010
Average	-47.9%	2.49	-47.3%	2.99	-32.6%	1.39	-35.2%	1.30	-46.8%	0.004

Next steps

- We have now entered the collaborative phase of the activity:
  - Refine rate distribution between geometry and colour to improve RD performance
  - Training JPEG AI on projected image data to improve colour coding performance
  - Implement residual lossless coding in geometry only pipeline to enable a lossless coding mode
  - Improving diversity of the test set



Next Steps

Examine recent developments in point cloud coding for inclusion in the VM:

- Testing with a sparse convolution framework to improve encoding/decoding times
- Attention models to improve performance by helping the model to focus on key aspects of the data – potential integration into the DL-based Block Encoding

Next Steps (my thoughts)

Transformers in point cloud applications show great promise:

- Detection
- Classification
- Registration
- Tracking
- Segmentation
- **Coding**

The inherent order invariance of Transformers is suited to point cloud data as well as the ability to capture global details via self-attention.

The application of transformers to many tasks matches well JPEG's goal of a latent representation that supports all of these functionalities.

Activity Timeline

Part	Title	WD	CD	DIS	FDIS	IS
6	JPEG Pleno: Learning-based Point Cloud Coding	23/01	23/10	24/04	-	25/01

How to participate:
JPEG Pleno Point Cloud Ad hoc Group

Chair: Stuart Perry (University of Technology Sydney, AU)
Co-Chair: Luis Cruz (University of Coimbra, PT)

- Ad hoc Group on JPEG Pleno - Lightfield
- Ad hoc Group on JPEG Pleno - Holography
- **Ad hoc Group on JPEG Pleno - Point Clouds**

Email reflector: **jpeg-Pointcloud**
To subscribe to the reflector, please visit http://listregistration.jpeg.org
or in case of problems contact lists@jpeg.org

Thank you!

Thanks to: Luis Cruz, Andre Guarda, Nuno Rodrigues, Fernando Pereira and Zhe Luo

To find out more:

https://jpeg.org/

Stuart.Perry@uts.edu.au