Studying the Benefits of a New JPEG AI Profile for the JPEG PCC Verification Model

Doc. WG1m100115

Abdelrahman Seleem, *André Guarda, Nuno Rodrigues, Fernando Pereira*
Instituto de Telecomunicações, Instituto Superior Técnico, and Instituto Politécnico de Leiria
PORTUGAL

July 2023

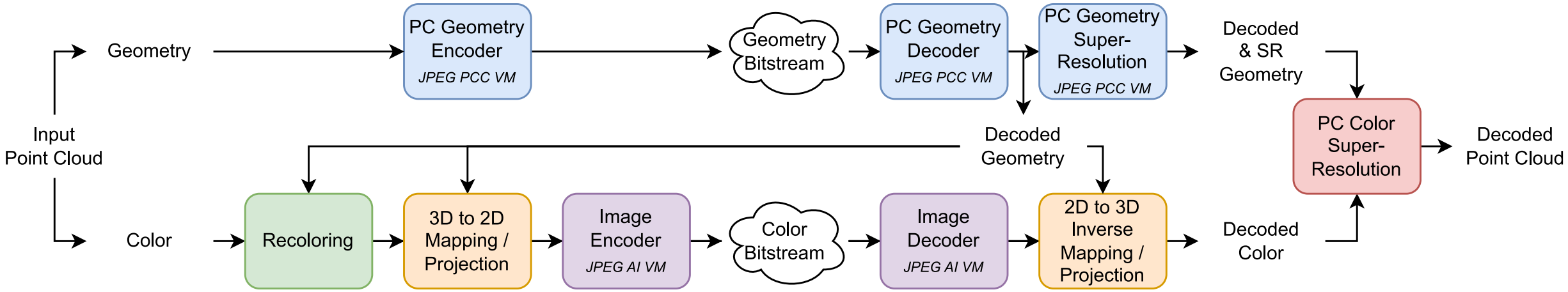
Outline

1. Objective
2. Retraining JPEG AI
3. Experimental Results
4. Conclusions and Next Steps

Context and Objective

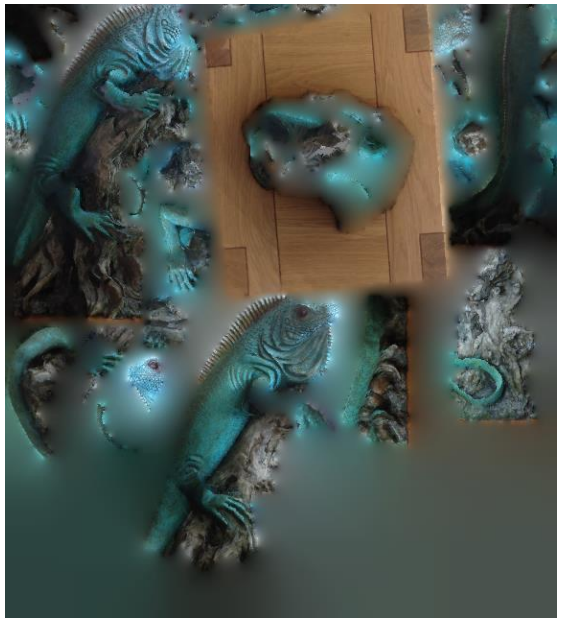
Context

- **The current JPEG PCC VM color coding approach first projects the PC color onto 2D images, then uses JPEG AI to code the 2D images.**



Motivation

- **The available JPEG AI models were trained for natural images.**
- **The 2D images created by the V-PCC 3D to 2D projection in the JPEG PCC VM do not really look like natural images.**



Objective

Improve the JPEG Pleno PCC VM color coding performance by retraining the available JPEG AI models with the images resulting from the projected point cloud color.

Retraining JPEG AI

JPEG AI Training Stages

1. **MSE_FixedRate_64:**
   - Loss function: MSE
   - Only 4 models and only 4 rates
2. **Mixed_FixedRate_36:**
   - Loss function: (MSE + MS-SSIM)/2
   - Only 4 models and only 4 rates
3. **Mixed_FixedRate_OnlyDec_20:**
   - Loss function: (MSE + MS-SSIM)/2
   - Only 4 models and only 4 rates
4. **MSE_VariableRate_12:**
   - Loss function: MSE
   - Only 4 models but multiple possible rates

Configurations to be Studied: Training Conditions

JPEG AI	4th stage	No retraining	The existing models are used as-is (i.e., trained with the JPEG AI training dataset)
		Fine tuning	The models are initialized with the existing weights (i.e., trained with the JPEG AI training dataset), and then trained/fine tuned with the projected PC images, under the same training conditions.
	1st stage	No retraining	The existing models are used as-is (i.e., trained with the JPEG AI training dataset)
		Fine tuning	The models are initialized with the existing weights (i.e., trained with the JPEG AI training dataset), and then trained/fine tuned with the projected PC images, under the same training conditions.
		Retraining from scratch	The models are randomly initialized, and then trained with the projected PC images, under the same training conditions.

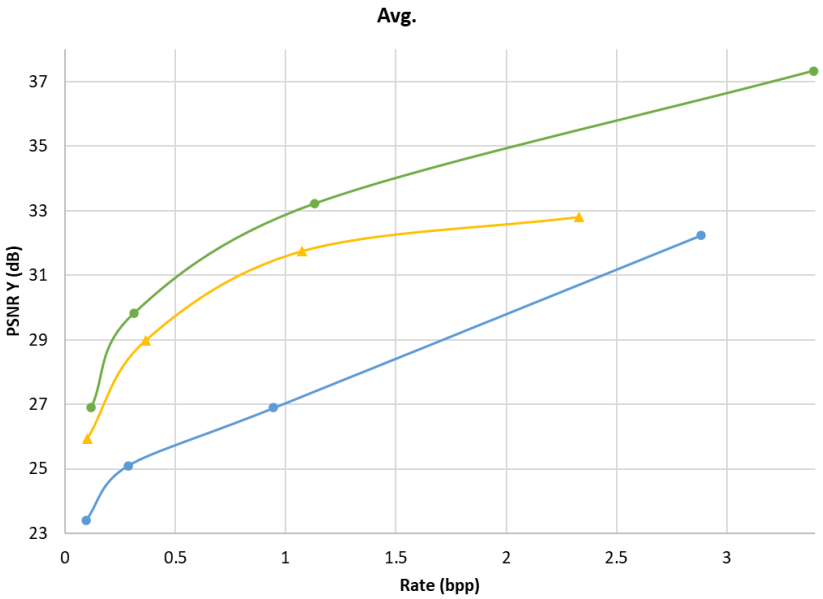
Training Dataset: Projected Images from JPEG Pleno PCC

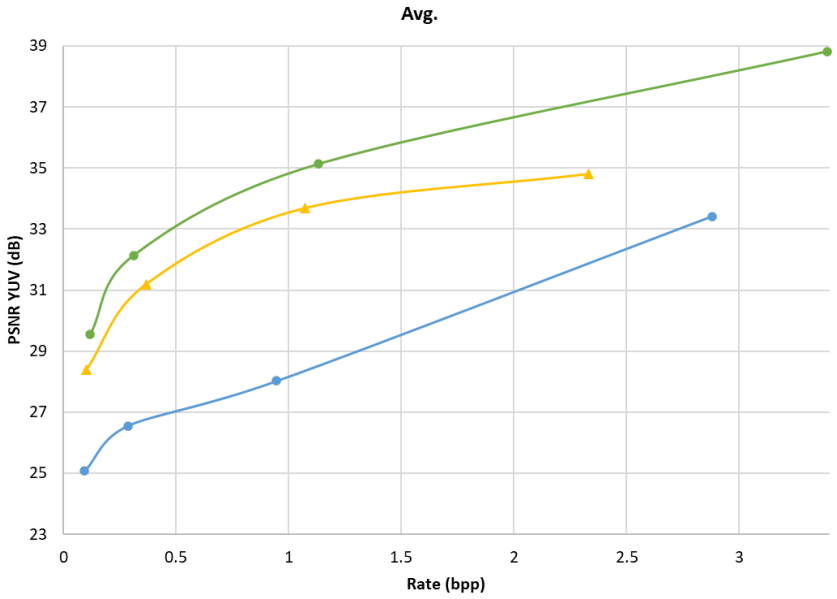
PC name	Sampling factor	Projected image resolution	# of Cropped images	
			Size: 1024x1024	
			Training	Validation
andrew10	1	1088x896	2	×
basketball_player_vox11_00000200	1	2112x2112	36	×
david10	1	1088x1088	16	×
Egyptian_mask_vox12	8	576x1344	2	×
exercise_vox11_00000001	1	2112x1728	24	×
Facade_00015_vox14	4	2112x2048	24	×
Facade_00064_vox14	16	1088x1024	2	×
football_vox11_1365600	2	1088x1088	16	×
Head_00039_vox12	2	2112x7104	84	×
Landscape_00014_vox14_color	16	1088x1216	16	×
loot_vox10_1200	1	1088x1472	16	×
mitch_vox11_00001	1	2112x2560	36	×
model_vox11_00000001	1	2112x1472	24	×
redandblack_vox10_1550	1	1088x1152	16	×
sarah10	1	1088x1088	16	×
Shiva_00035_vox12	4	1088x3392	32	×
soldier_vox10_0690	1	1088x1280	16	×
Stanford_Area_2_vox16_color	16	4160x5952	120	×
Stanford_Area_4_vox16_color	32	2112x4544	60	×
Thaidancer_viewdep_vox12	2	2112x1088	24	×
the20smaria_00600_vox11	1	2112x1856	24	×
thomassenic_vox11_00170	1	2112x2112	36	×
ULB_Unicorn_vox13_n	8	1088x3008	24	×
ulliwegner_01400_vox11	1	1088x1024	2	×
boxer_viewdep_vox12	2	2112x2432	×	36
dancer_vox11_00000001	1	2112x1984	×	24
Facade_00009_vox12	4	1088x1152	×	16
Frog_00067_vox12	2	2112x4608	×	60
Sum			**668**	**136**

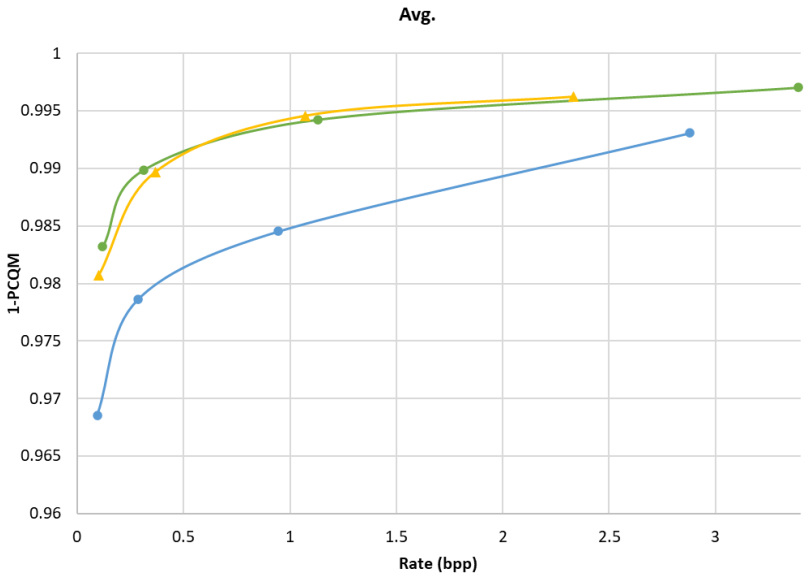
- ✶ Random cropping was used

Experimental Results

RD Performance: Using JPEG AI 4th Stage

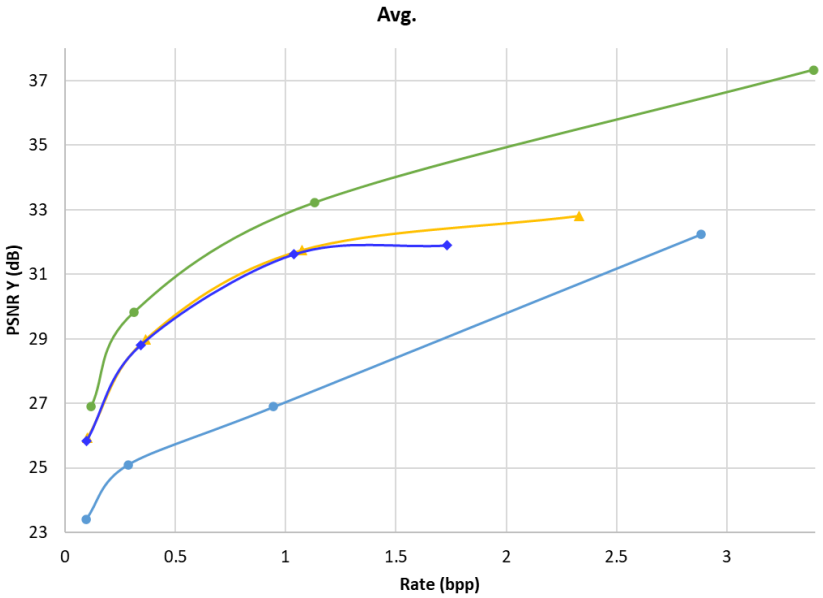


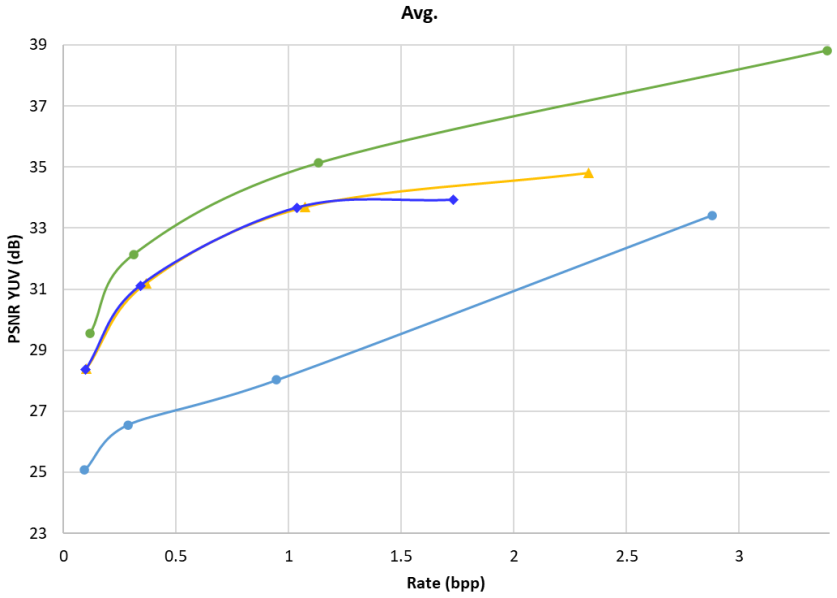


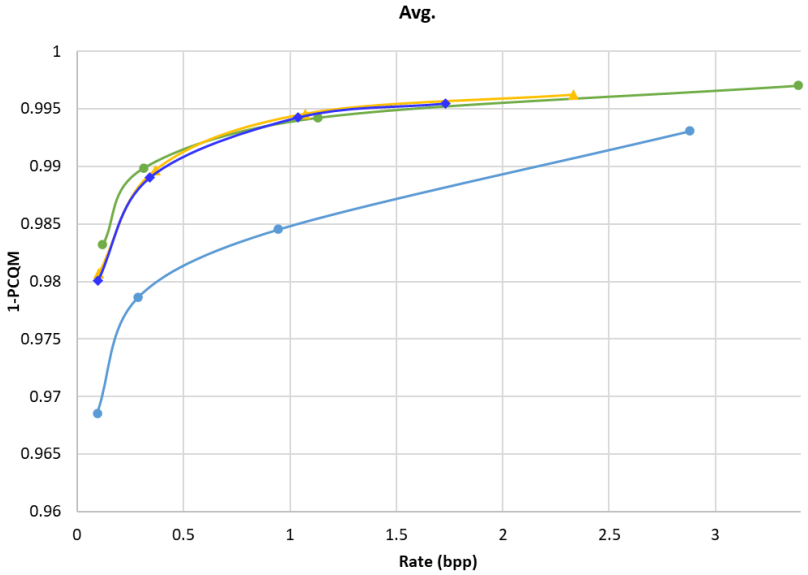


- JPEG PCC VM v2 (JPEG AI; 4th Stage no retraining) is the current VM

RD Performance: Using JPEG AI 1st Stage



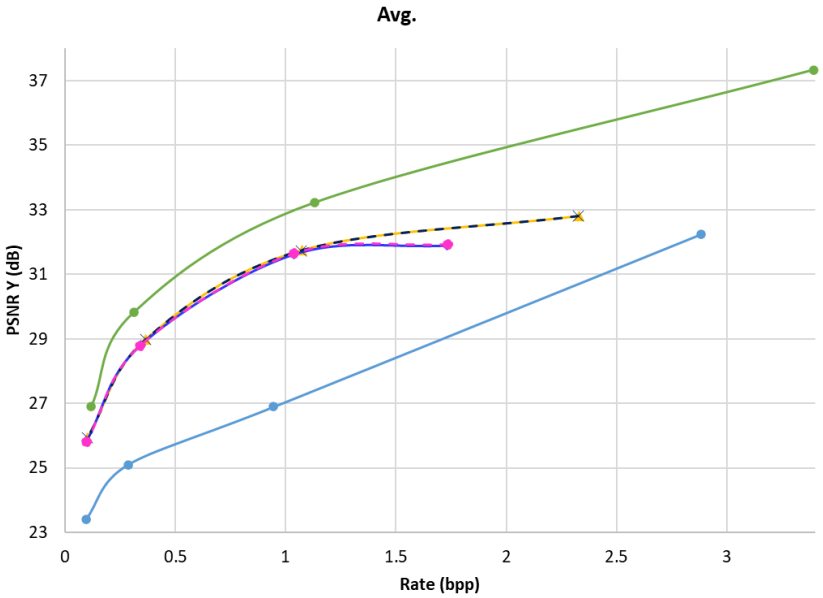


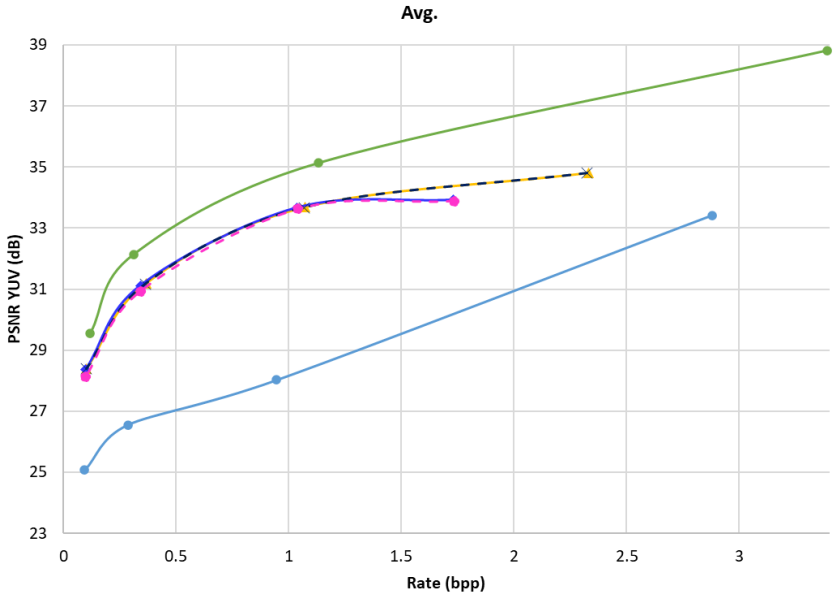


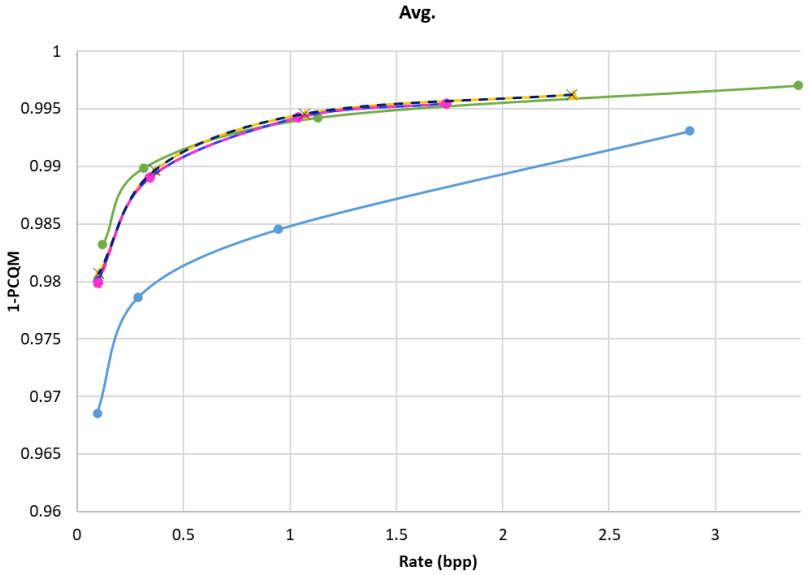
G-PCC Octree + PredLift v20
V-PCC Intra v20
JPEG PCC VM v2 (JPEG AI; 4th Stage no retraining)
JPEG PCC VM v2 (JPEG AI; 1st Stage no retraining)

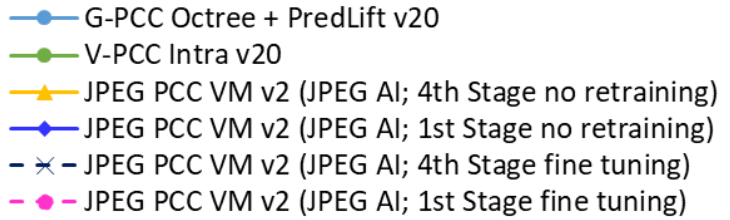
- Using the JPEG AI models from the 1st stage produces similar results to the models from the 4th stage, except for the last rate point since the 1st stage only considers a fixed rate

RD Performance: Fine Tuning JPEG AI



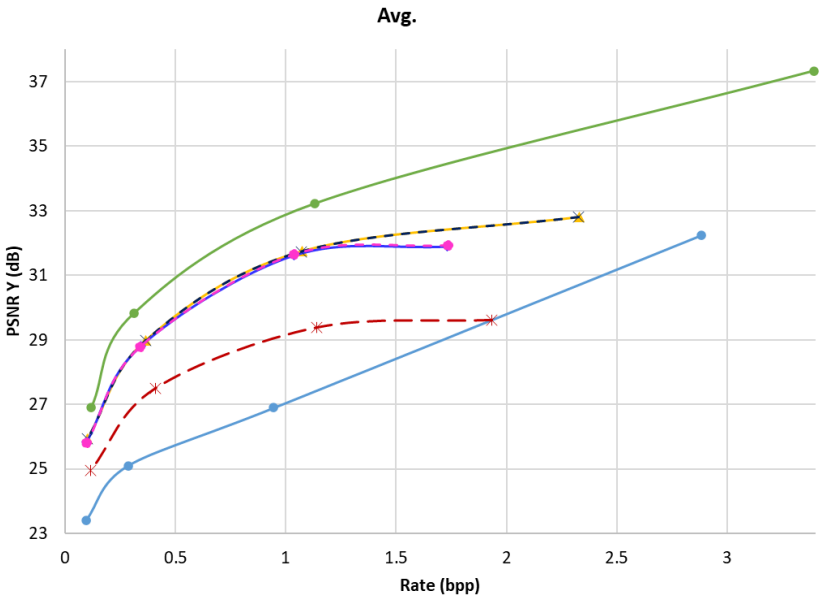


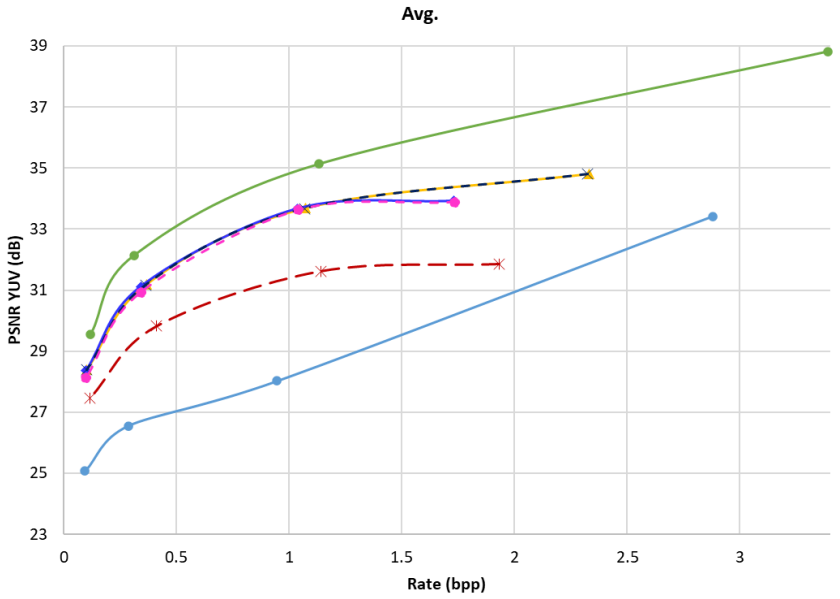


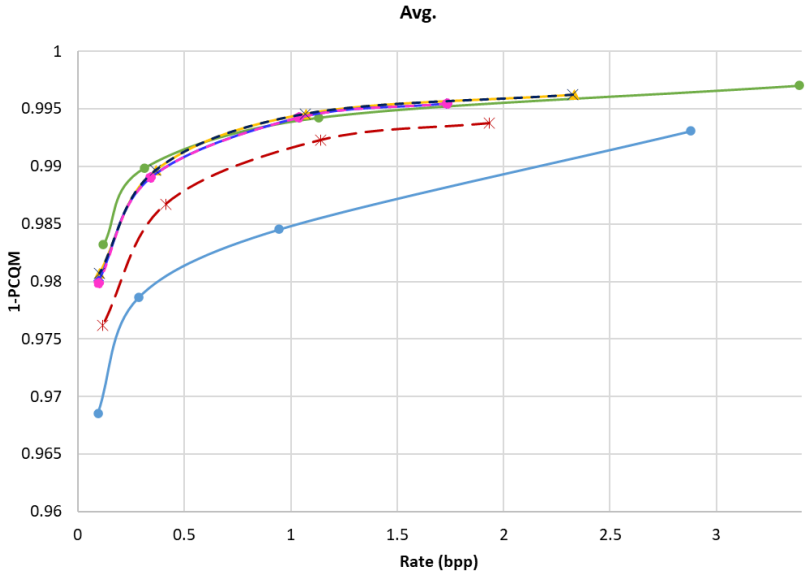


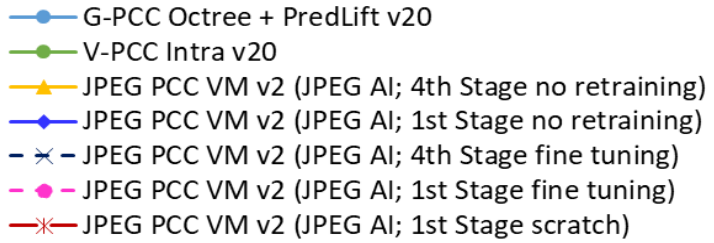
- Overall, fine tuning did not impact the results for neither the 1st or 4th stages

RD Performance: Fully Retraining JPEG AI 1st Stage









- Training the JPEG AI models from scratch (random initialization) severely downgrades the performance

Conclusions and Next Steps

Conclusions and Next Steps

- **The performed experiments of retraining the JPEG AI models using projected point cloud images did not demonstrate any improvement**
- **It is possible that the projected images training dataset is too small, not allowing the JPEG AI models to properly learn new information**
- **Further investigation is necessary, starting with the increase of the training data**