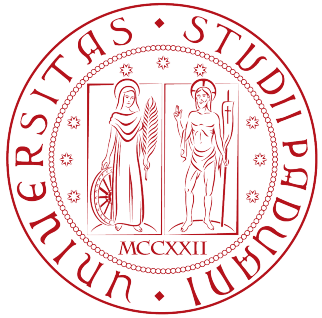
UNIVERSITY OF PADOVA



ENGINEERING SCHOOL

Master thesis:

Color coding for 3D models in augmented reality applications

Supervisor: Prof. Simone Milani

Candidate: Enrico Polo

Academic Year: 2016/17

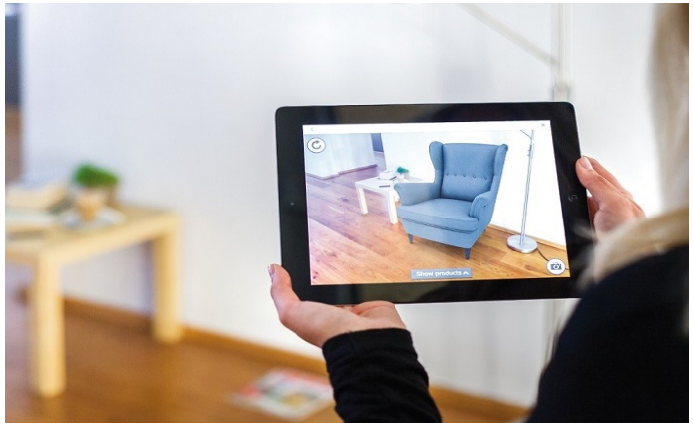
DEPARTMENT OF INFORMATION ENGINEERING
UNIVERSITY OF PADOVA

11 December 2017

Introduction

- What is augmented reality (AR)?

  - Human perception enhancement through computer-generated contents



- Which AR applications are we interested in?

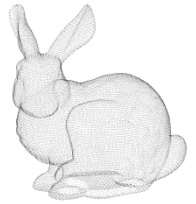
  -Visual AR

  -Real time applications

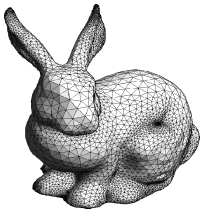
3D objects representation

- Requirements
  - High compression
  - Low computation
  - Good rendering

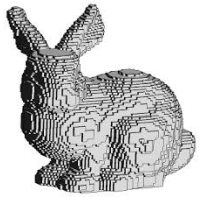
Point cloud



Mesh



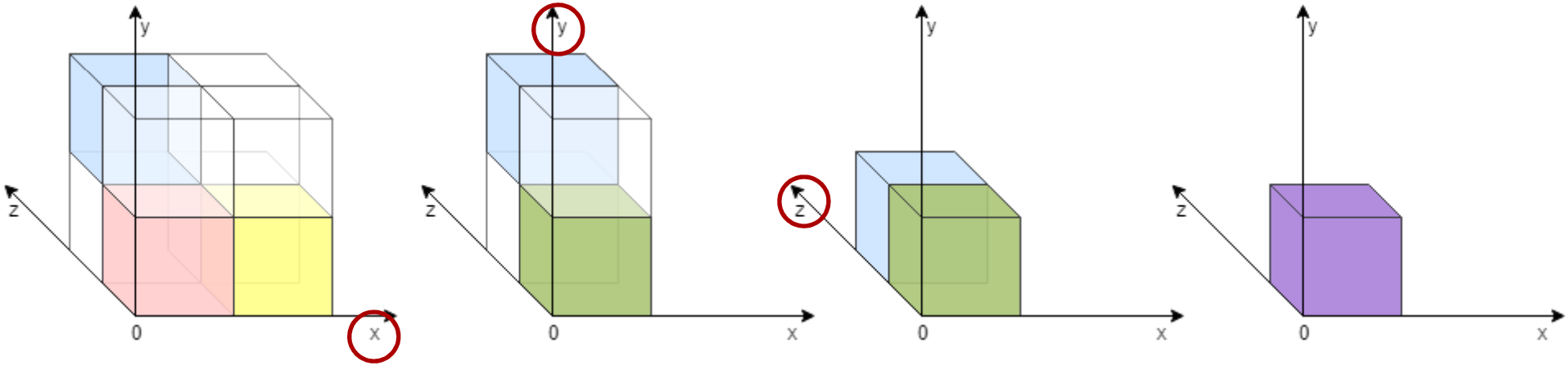
Voxel grid



- Goals
  - Color coder implementation for static objects
  - Color prediction scheme development for 3D videos

Region-Adaptive Hierarchical Transform

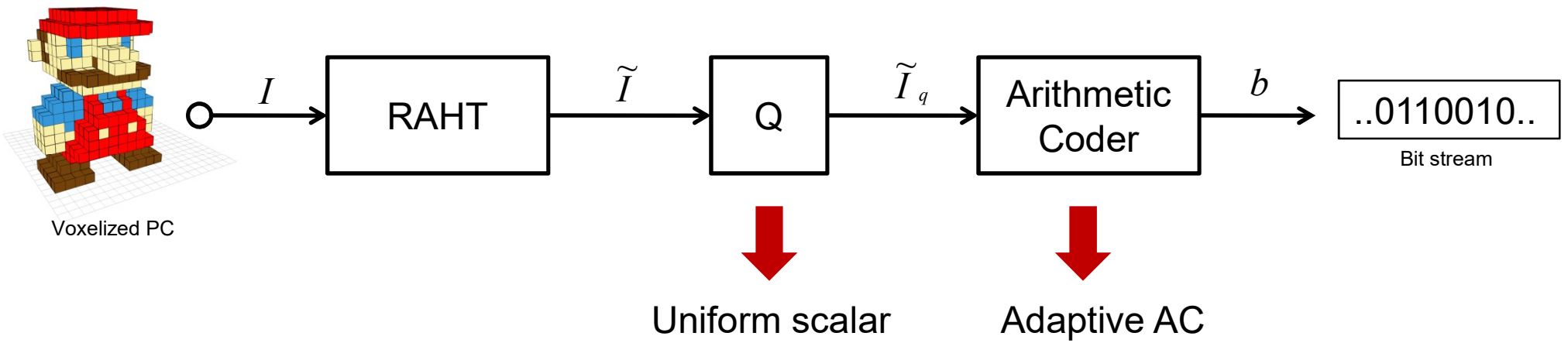
- Transform developed by Microsoft to compress voxelized attributes



- $\begin{pmatrix} DC \\ AC \end{pmatrix} = \frac{1}{\sqrt{w_1 + w_2}} \begin{pmatrix} \sqrt{w_1} & \sqrt{w_2} \\ -\sqrt{w_2} & \sqrt{w_1} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

- $w_1 = w_1 + 1$

Single frame (intra coding)



- RAHT v1

  - Color organized as a 3D structure

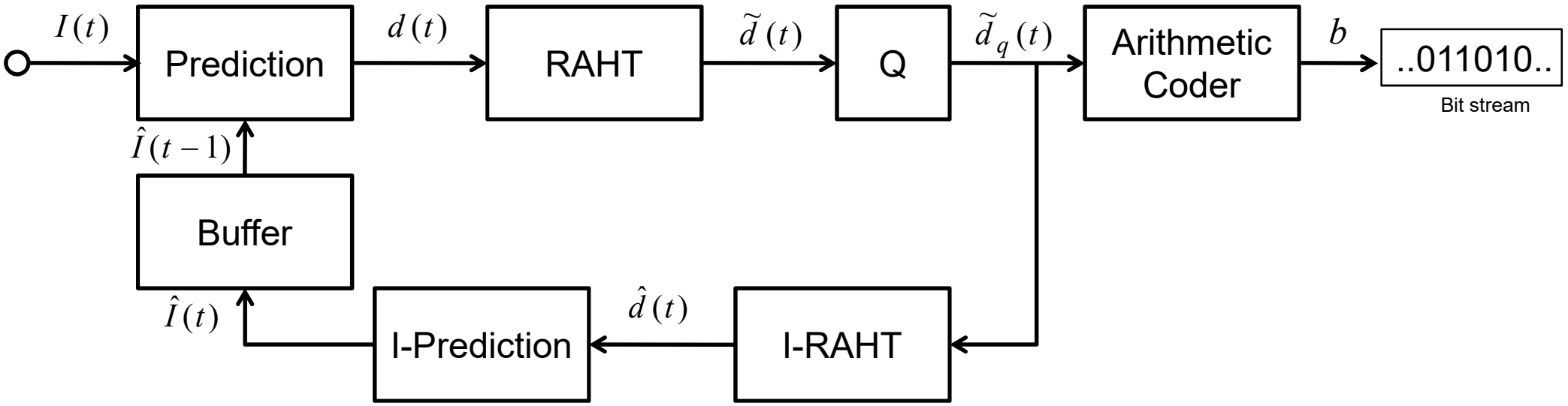
  - Requires full volume scansion

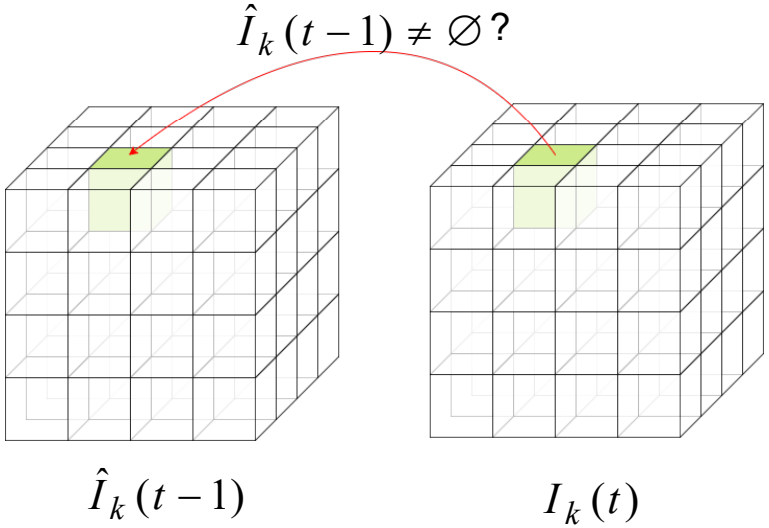
- RAHT v2

  - Color organized as an array of (color, coordinate) pairs

  - Requires coordinate update and sorting after each iteration

Prediction (inter coding)





- Yes: $d_k(t) = I_k(t) - \hat{I}_k(t-1)$
- No: $d_k(t) = I_k(t) - \overline{\hat{I}_k(t-1)}$

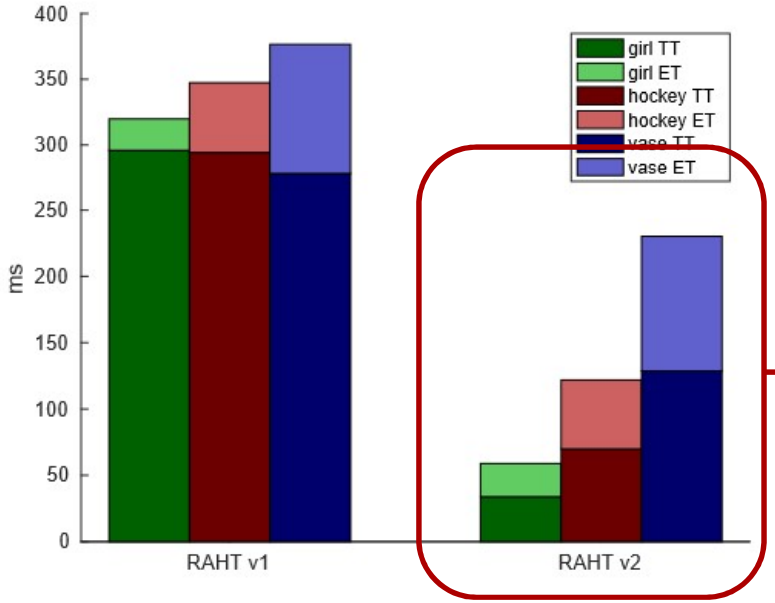
Results: intra coding complexity

- Girl: 0.1% o.v.

- Hockey: 0.2% o.v.

- Vase: 0.4% o.v.

- Resolution: 512^3



RAHT v2

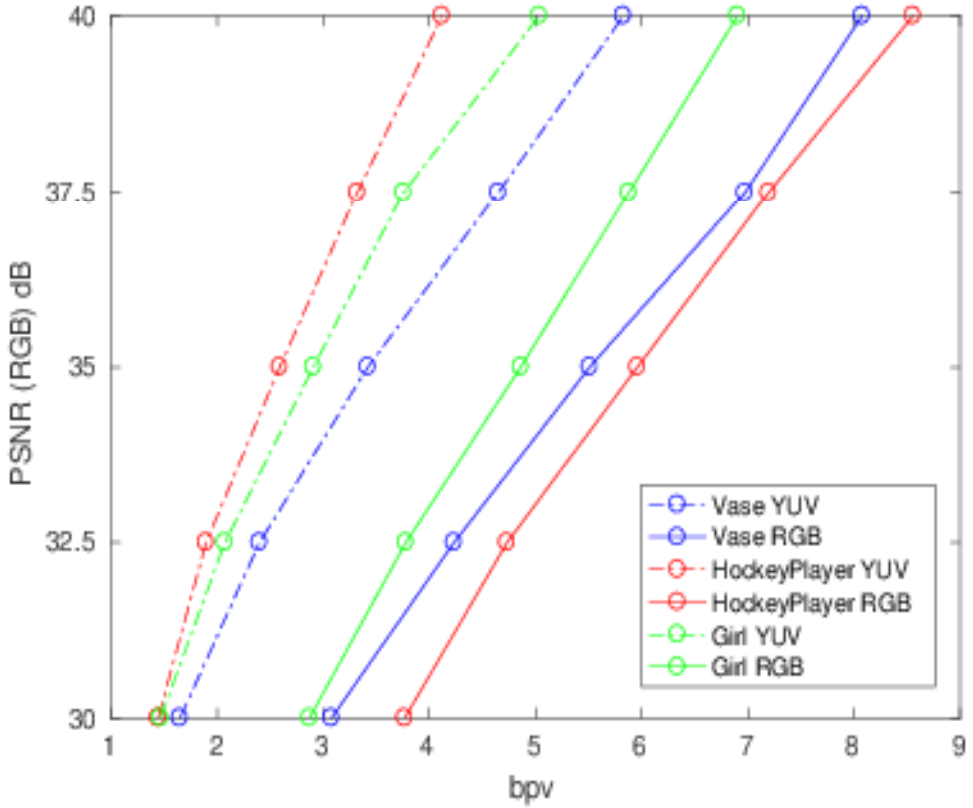
Results: intra coding compression

- Distortion metric:

$$PSNR = 20 \log_{10}\left(\frac{255}{\sqrt{MSE}}\right) \quad [dB]$$

- Rate metric:

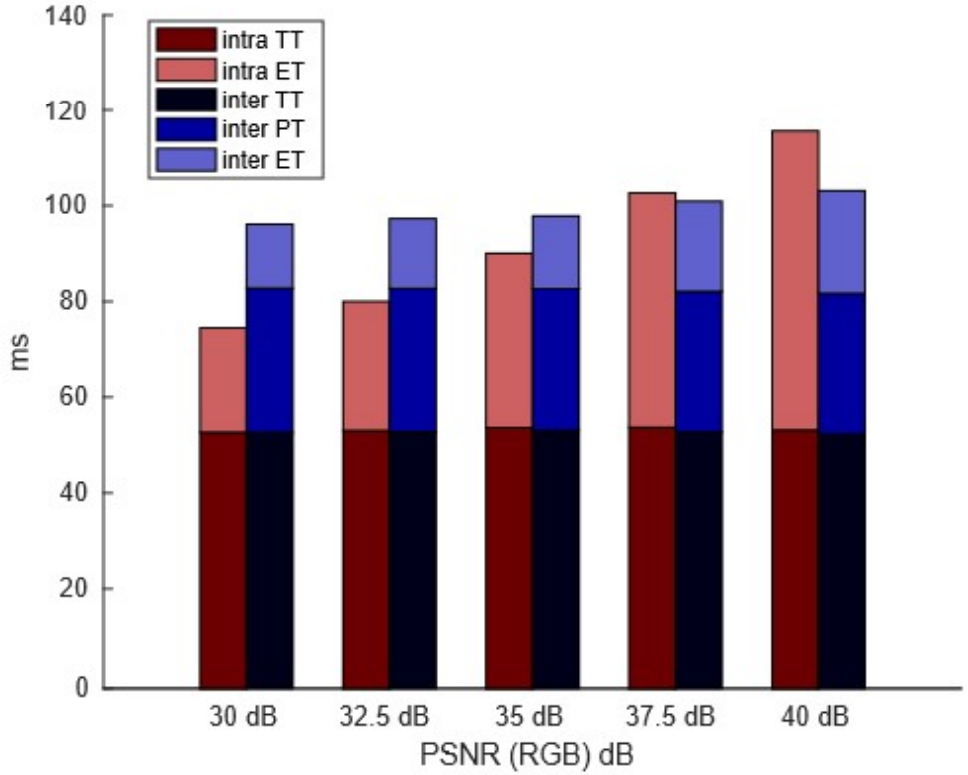
$$bpv = \frac{B_{tot}}{N_{occ}}$$



- YUV color space must be preferred

Results: inter coding complexity

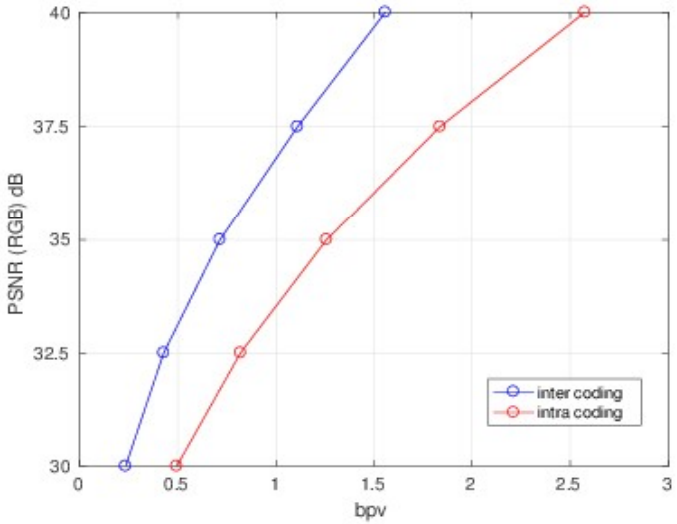
- Resolution: 512^3
- Sarah: 0,2% avg. o.v.



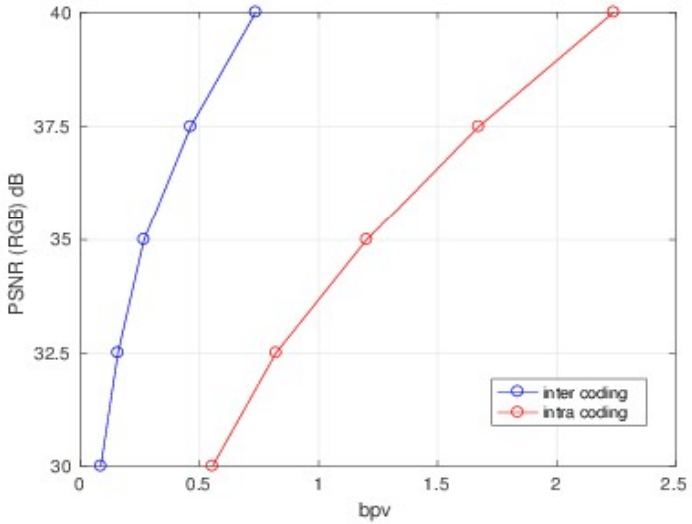
- Frames: 207

"Voxelized upper bodies dataset." Available: https://jpeg.org/plenodb/ pc/microsoft/

Results: intra vs inter compression



Ricardo



Sarah

Sequence	30dB	32.5dB	35dB	37.5dB	40dB	Average
Andrew	69.09%	64.30%	54.18%	49.09%	49.80%	57.29%
David	32.22%	26.70%	21.55%	19.05%	18.30%	23.56%
Phil	39.65%	34.51%	30.64%	29.02%	29.08%	32.58%
Ricardo	52.81%	47.80%	43.19%	39.56%	39.52%	44.57%
Sarah	84.88%	81.03%	77.93%	72.39%	67.19%	76.68%

Results: Microsoft comparison

- Milani-Limuti-Polo:

  - Lossless geometry transform (intra coding)

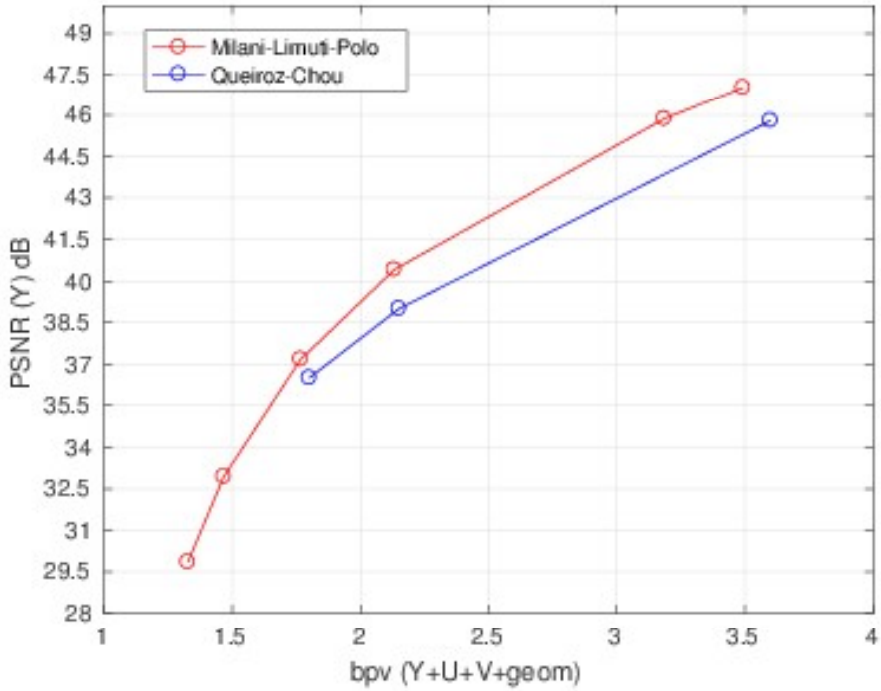
  - RAHT color transform (inter coding)

- Queiroz-Chou:

  - Lossy geometry coding

  - RAHT color transform

  - Motion-compensated prediction for both



R. L. de Queiroz and P. A. Chou, “Motion-compensated compression of dynamic voxelized point clouds,” IEEE Transactions on Image Processing, vol. 26, no. 8, pp. 3886–3895, Aug 2017

Conclusion

- Execution time:

  - For resolutions < 512^3 we satisfy real-time constraints (30 fps)

  - For 512^3 resolution we are able to encode 6-10 fps

  → Parallelization

  → GPU programming

- Compression:

  - Prediction allows us to save (on average) about 50% of the bits

  - Effective performances depends on the specific subject

  → AC optimization