PhD Proposal

3D Point Cloud Compression using Deep Neural Networks

Key words: 3D point clouds, deep learning, auto-encoders, octrees, voxels, quality

Location: InterDigital Rennes / CentraleSupelec Paris

Supervisors: Jean-Eudes MARVIE / Giuseppe VALENZISE, Frédéric DUFAUX

Summary: Three-dimensional point clouds are set of points in the 3D space, associated to attributes such as colors, reflectance, normals, etc. They are an essential data structure in several domains, including virtual and mixed reality, immersive communication and perception in autonomous vehicles. Since point clouds easily range in the millions of points and can have complex sets of attributes, efficient point cloud compression (PCC) is fundamental to enable the consumption of point cloud content in practical applications. As a result, compression of point clouds is currently a matter of research and standardization, e.g., the Moving Picture Expert Group (MPEG) has launched a standardization initiative to compress the geometry and/or attributes of static or dynamic point clouds [1]. This PhD topic concerns the study and implementation of new solutions based on deep learning in order to improve existing point cloud compression methods. Recent results in this area [2,3,4] are very promising in terms of compression performance compared to traditional approaches. However, they still have many limitations, especially regarding the nature and volume of input data. Indeed, no existing deep learning-based solution allows the coding of geometry (spatial coordinates of points) and photometry (colors of these points, normals, etc.) in a unified and efficient way. Thus, our objective consists in bringing together these different approaches in order to obtain viable solutions in terms of industrial use. This convergence will be obtained in particular through the definition of new representations of input geometric and photometric data allowing consumption in regular data structures compatible with deep learning models, as well as by the definition of new perceptual loss functions yielding better compression gains and high visual quality reconstruction. In terms of data structures, existing deep-learning-based solutions are generally applied to regular voxelizations of the point cloud. This type of representation introduces strong degradations on the original model (quantization, tile discontinuities, etc.) and provides poor scalability when the size of the point cloud increases. We therefore propose to enhance the existing algorithms in order to make the best use of multi-resolution data structures (e.g., octrees). This enables scalability via progressive refinement, as well as the view dependent consumption of the data stream. Here again, an adaptation of the context data, from the descending tree traversal, must be performed in order to be supplied to the learning model. Finally, we propose in a last phase to study the possibility of extending our model to animated point clouds, by considering the temporal aspect and therefore inter-frame compression. Following an in-depth state of the art, the student will therefore have to develop new perceptual metrics, data structures and algorithms allowing the implementation of these different aspects.

Position details: The PhD thesis will be partially funded by the ANRT CIFRE funding scheme, which involves the collaboration of a research lab (L2S, CentraleSupelec, University Paris-Saclay) and an industrial partner (InterDigital, Rennes). The PhD student will be hired by Interdigital and enrolled to the STIC (Sciences et Technologies de l'Information et de la Communication) doctoral school of the university Paris-Saclay. The position is supposed to start in early 2021.

Research labs: The Point Cloud Compression team of InterDigital Imaging Science R&D laboratory has the mission is to understand the industrial needs in the field of point cloud compression, and to

propose innovative solutions to the MPEG PCC standard, federated around complete demonstrators. The student will benefit from an existing software infrastructure which will allow him to concentrate on the research / optimization part of technologies and to minimize the software developments to be carried out. A large hardware infrastructure of many core CPUs and GPUs will be at his disposal for model training. In addition, he will have the opportunity to exploit the scientific expertise concentrated on the same site, particularly in terms of video compression, point cloud compression and streaming of large 3D environments [5, 6, 7].

The Laboratoire des Signaux et Systèmes (L2S) is a joint research unit (UMR 8506) of the "Centre National pour la Recherche Scientifique" (CNRS), CentraleSupélec and the Paris-Saclay University. The research of L2S is about the fundamental and applied mathematical aspects of control theory, signal and image processing, information and communication theory. The student will integrate the "Reseau Optimisation Conjointe" (ROC) team in the Telecommunication&Networking division and will benefit from the experience of researchers in the team on video compression, quality assessment and deep learning. In particular, the advisors, G. Valenzise and F. Dufaux, have a long experience in video compression, including coding of 3D and immersive media [9], applications of deep learning to enhance video compression performance [8], as well as subjective and objective methodologies to assess the visual quality of experience [10, 11, 12]. In addition, the team in L2S has pioneered the use of deep learning approaches for compressing point clouds [3, 4].

Required skills: Prospective PhD candidates are expected to have a strong background both in signal processing, applied math, and programming, in particular with Python (possibly with deep learning frameworks such as TensorFlow or Pytorch) and C++ (the reference code of MPEG G-PCC is written in this language). Fluency in English and good communication skills in general are highly required.

Contacts: Send your application with a CV to Jean-Eudes.Marvie@interdigital.com, giuseppe.valenzise@l2s.centralesupelec.fr and frederic.dufaux@l2s.centralesupelec.fr

References:

[1] S. Schwarz et al. "Emerging MPEG standards for point cloud compression." IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 9, no. 1, March 2019.

[2] A. F. R. Guarda, N. M. M. Rodrigues, and F. Pereira, "Deep learning-based point cloud coding: A behavior and performance study," in 2019 8th European Workshop on Visual Information Processing (EUVIP), pp.34–39, ISSN: 2164-974X

[3] M. Quach, **G. Valenzise**, and **F. Dufaux**, "Learning convolutional transforms for lossy point cloud geometry compression," in IEEE International Conference on Image Processing (ICIP), pp. 4320–4324, Taipei, Taiwan, September 2019

[4] M. Quach, **G. Valenzise** and **F. Dufaux**, "Folding-based compression of point cloud attributes," in IEEE International Conference on Image Processing (ICIP), 2020

[5] P. Gautron, C. Delalandre, **J.-E. Marvie** and Pascal Lecocq. "Boundary-Aware Extinction Mapping." In proceedings of *Pacific Graphics 2013*. October 2013. Singapor.

[6] **J.-E. Marvie**, G. Pascal, P. Lecocq, O. Mocquard and F. Gerard. "Streaming and Synchronization of Multi-User Worlds Through HTTP/1.1." In proceedings of ACM Siggraph *Web3D 2011*. June 2011. Paris, France.

[7] R. Lerbour, **J.-E. Marvie** and P. Gautron, "Adaptive Streaming and Rendering of Large Terrains: A Generic Solution**."** In proceedings of the *17th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*. February 2009. Plzen - Bory, Czech Republic.

[8] L. Wang, A. Fiandrotti, A. Purica, **G. Valenzise** and M. Cagnazzo, "Enhancing HEVC Spatial Prediction by Context-based Learning," ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, United Kingdom, 2019, pp. 4035-4039, doi: 10.1109/ICASSP.2019.8683624.

[9] **F. Dufaux**, B. Pesquet-Popescu, and M. Cagnazzo, "Emerging technologies for 3D video: creation, coding, transmission and rendering." John Wiley & Sons, 2013

[10] N. K. Kottayil, **G. Valenzise**, **F. Dufaux** and I. Cheng, "Blind Quality Estimation by Disentangling Perceptual and Noisy Features in High Dynamic Range Images," in IEEE Transactions on Image Processing, vol. 27, no. 3, pp. 1512-1525, March 2018, doi: 10.1109/TIP.2017.2778570.

[11] M. Pérez-Ortiz, A. Mikhailiuk, E. Zerman, V. Hulusic, **G. Valenzise** and R. K. Mantiuk, "From Pairwise Comparisons and Rating to a Unified Quality Scale," in IEEE Transactions on Image Processing, vol. 29, pp. 1139-1151, 2020, doi: 10.1109/TIP.2019.2936103.

[12] **G. Valenzise**, A. Purica, V. Hulusic, M. Cagnazzo, "Quality assessment of deep-learning-based image compression", in Proc. IEEE Multimedia Signal Processing Workshop, Vancouver, CA, August 2018