



HAL
open science

Compression de Nuages de Points basée sur Normalizing Flow

Rodrigo Borba Pinheiro, Jean-Eudes Marvie, Giuseppe Valenzise, Frédéric Dufaux

► **To cite this version:**

Rodrigo Borba Pinheiro, Jean-Eudes Marvie, Giuseppe Valenzise, Frédéric Dufaux. Compression de Nuages de Points basée sur Normalizing Flow. Colloque COMpression et REprésentation des Signaux Audiovisuels (CORESA'2024), Nov 2024, Rennes, France. hal-04826722

HAL Id: hal-04826722

<https://hal.science/hal-04826722v1>

Submitted on 12 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Compression de Nuages de Points basée sur Normalizing Flow

R.B. Pinheiro^{1,2}

J.E. Marvie¹

G. Valenzise²

F. Dufaux²

¹ InterDigital, Inc.

² Université Paris-Saclay, CNRS, CentraleSupélec, L2S 91190 Gif-sur-Yvette, France

{rodrigo.borbapinho, jean-eudes.marvie}@interdigital.com

{giuseppe.valenzise, frederic.dufaux}@centralesupelec.fr

Résumé

Les méthodes existantes basées sur l'apprentissage profond pour compresser les attributs des Nuages de Points (NdP) utilisent généralement des autoencodeurs variationnels (VAE). Toutefois, ces schémas souffrent d'une qualité de reconstruction limitée aux débits élevés. Nous proposons une architecture basée sur Normalizing Flow (NF), appelée NF-PCAC, comme alternative. Les NFs sont des réseaux inversibles qui permettent une reconstruction de haute qualité. Nous proposons aussi une amélioration de l'architecture avec une complexité réduite, appelée RNF-PCAC. Elle se compose de deux modes de fonctionnement spécialisés : un pour les bas débits et un pour les hauts débits. Notre approche réduit le nombre de paramètres de NF-PCAC de plus de 6×. Nos architectures surpassent les méthodes précédentes basées sur l'apprentissage pour la compression d'attributs et nous atteignons des performances comparables à celles de G-PCC (v.21)

Mots clefs

Compression, Nuages de points, Apprentissage profond, Réseaux de neurones, Normalizing Flow

1 Introduction

La consommation de contenu a évolué vers des formats immersifs, en particulier pour le divertissement. Dans ce contexte, les nuages de points (NdP) sont l'une des représentations volumétriques les plus populaires. Les NdPs sont des ensembles de points non ordonnés qui contiennent les coordonnées x, y, z et les informations d'attribut respectives, dans notre cas, les couleurs rouge, vert, bleu. Cependant, pour représenter un objet, les NdPs doivent contenir des millions de points, ce qui devient coûteux à stocker et à transmettre. La compression est donc nécessaire pour en faire une option viable pour la diffusion de contenu 3D. Les efforts du groupe MPEG [1] ont contribué de manière significative aux avancées dans le domaine de la compression de NdP et, plus récemment, les techniques basées sur l'apprentissage ont gagné en popularité.

Deep-PCAC [2] utilise des convolutions par point de second ordre, qui sont plus légères, mais qui sont incapables de capturer les dépendances spatiales. Sparse Tensor PCAC [3] utilise une architecture VAE et des convolutions

sparses. Les VAE sont efficaces à bas débit, car ils réduisent la dimensionnalité des données. Toutefois, cette réduction limite la qualité de reconstruction à haut débit. Nous proposons des alternatives basées sur NF pour coder les attributs des NdP : «Normalizing Flow Point Cloud Attribute Compression» (NF-PCAC) et «Reduced Normalizing Flow Point Cloud Attribute Compression» (RNF-PCAC).

2 Notre approche

NF est un réseau qui modélise des transformées inversibles et peut, en principe, réaliser une reconstruction sans perte. NF-PCAC adapte une architecture 2D pour prendre en compte la nature 3D et l'aspect diffus des NdPs. Pour augmenter le champ réceptif des convolutions, nous avons conçu une couche appelée "3D Squeeze Layer" [4], qui remplace les voxels vides par la moyenne de leur voisinage. De cette façon, nous améliorons la convergence de l'algorithme et obtenons de meilleurs résultats. Pour réduire la complexité, nous avons par la suite conçu RNF-PCAC [5], un réseau composé de deux modes de fonctionnement. Le mode Back Projection (BP) qui cible la plage de bas débit et le mode No-Average (NA) que cible celle de haut débit. Nous choisissons dynamiquement le meilleur mode en utilisant une approche d'optimisation de débit-distorsion. Notre objectif est de contrôler le nombre de paramètres dans le réseau tout en obtenant de bonnes performances. Nous obtenons de meilleurs résultats que NF-PCAC, en réduisant de six fois le nombre de paramètres, et nous obtenons également des résultats comparables à G-PCC version v.21 sur certains NdPs.

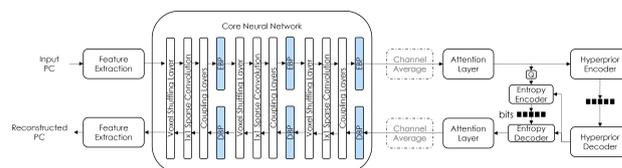


FIGURE 1 – **Mode BP** : Les couches de Back Projection de l'encodeur (EBP) et les couche de Back Projection du décodeur (DBP) nous permettent de contrôler le nombre de canaux à travers le réseau et à réduire le nombre de coefficients du réseau. Blocs ajoutés en bleu et blocs supprimés en gris pointillés, par rapport à NF-PCAC [4]

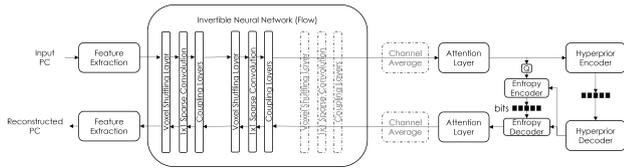
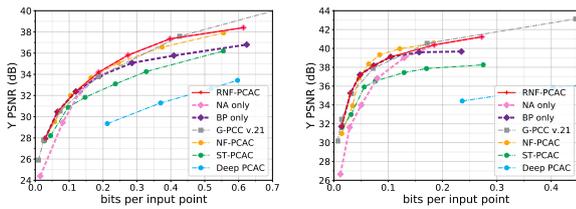


FIGURE 2 – **Mode NA** : Les sources de non-inversibilité sont réduites. Nous obtenons une version plus petite du réseau initial avec un espace latent plus grand. Blocs supprimés en gris pointillés par rapport à NF-PCAC [4].

3 Résultats

Comme le montre la Figure 3, notre première méthode NF-PCAC surpasse les autres approches basées sur l'apprentissage, mais avec un grand nombre de paramètres (voir Tableau 1). Le mode BP de RNF-PCAC surpasse toutes les autres méthodes sur la plage de bas débits, tout en utilisant 10 fois moins de paramètres que NF-PCAC pendant que le mode NA produit des gains de codage plus importants à des débits plus élevés, toujours en utilisant plus de 10 fois moins de paramètres que NF-PCAC (voir Figure 3 et Tableau 1).



(a) Soldier - 8iVFB



(b) Basketball - Owlfi

FIGURE 3 – Courbes PSNR pour deux NdPs différents. RNF-PCAC obtient des résultats compétitifs et surpasse d'autres méthodes basées sur l'apprentissage.

Méthodes	Params	Stockage	Peak GPU Memory				Temps			
			10-bits		11-bits		10-bits		11-bits	
			Enc	Dec	Enc	Dec	Enc	Dec	Enc	Dec
NF-PCAC	278	1.10	4.99	4.39	10.02	10.04	0.99	0.95	3.37	3.27
NA only	26	0.10	3.31	2.74	8.48	9.03	0.65	0.56	2.20	2.02
BP only	18	0.07	2.79	2.19	7.19	7.36	0.49	0.43	1.59	1.56
RNF-PCAC	44	0.17	3.31	2.74	8.48	9.03	1.14	0.56	3.78	2.02
ST-PCAC	12	0.05	1.85	1.08	3.82	3.47	0.11	0.07	0.33	0.22

TABLEAU 1 – Comparaison de la complexité : Millions de paramètres, stockage en Go, mémoire maximale utilisée en Go et temps moyen en secondes sur GPU (Nvidia®. Tesla®. V100, 32 Go). Résultats ST-PCAC à titre informatif.

Nous présentons les résultats du BD-Rate dans le tableau 2, en utilisant G-PCC comme référence. Les méthodes que

nous proposons obtiennent le meilleur résultat parmi les approches basées sur l'apprentissage.

Point Cloud	Basketball	Soldier	Longdress	Facade
Deep-PCAC	942.75	295.44	340.23	526.74
ST-PCAC	73.18	43.53	53.79	168.41
NF-PCAC (Ours) [4]	13.47	1.26	30.58	55.66
RNF-PCAC (Ours) [5]	-0.11	-4.80	17.99	44.00

TABLEAU 2 – Le BD-Rate (%) moyen pour le canal luma sur 4 NdPs différents : Méthodes basées sur apprentissage profond versus G-PCC v.21. Les valeurs négatives correspondent à des gains en débit.

4 Conclusion

Nous proposons NF-PCAC, la première architecture basée sur NF comme outil de compression des attributs des nuages de points. Nous avons amélioré cette architecture avec RNF-PCAC, une méthode comprenant deux modes de codage : BP Mode et NA mode. Ces deux modes contribuent à réduire l'empreinte mémoire et la complexité du codage des attributs NdP basé sur les NF, tout en augmentant les performances de codage. Nos résultats montrent que les approches basées sur la NF sont une voie prometteuse pour combler l'écart de performance entre les méthodes basées sur l'apprentissage et le G-PCC, offrant une alternative viable aux approches basées VAE.

Annexe

Les travaux ont déjà été publiés dans ICASSP 2023 [4] et ICASSP 2024 [5].

Références

- [1] D. Graziosi, O. Nakagami, S. Kuma, A. Zaghetto, T. Suzuki, et A. Tabatabai. An overview of ongoing point cloud compression standardization activities : video-based (V-PCC) and geometry-based (G-PCC). *APSIPA Trans. Signal and Inf. Process.*, 2020.
- [2] X. Sheng, L. Li, D. Liu, Z. Xiong, Z. Li, et F. Wu. Deep-PCAC : An end-to-end deep lossy compression framework for point cloud attributes. *IEEE Transactions on Multimedia*, 2021.
- [3] J. Wang et Z. Ma. Sparse tensor-based point cloud attribute compression. Dans *IEEE Int. Conf. on Multimedia Information Processing and Retrieval*, 2022.
- [4] R. B. Pinheiro, J.E. Marvie, G. Valenzise, et F. Dufaux. Nf-pcac : Normalizing flow based point cloud attribute compression. Dans *ICASSP 2023 - 2023 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5, 2023.
- [5] R.B. Pinheiro, J.E. Marvie, G. Valenzise, et F. Dufaux. Reducing the complexity of normalizing flow architectures for point cloud attribute compression. Dans *ICASSP 2024 - 2024 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8170–8174, 2024.