Evaluation de la qualité sans référence des nuages de points basée sur les statistiques de co-occurrence 3D

S. Riache et M.-C. Larabi CNRS, XLIM UMR 7252, Université de Poitiers Poitiers, France

Résumé

L'évaluation de la qualité des nuages de points reste un défi majeur en raison de la complexité des applications associées et de la nature du contenu. Pour résoudre ce problème, cet article propose une nouvelle métrique d'évaluation de la qualité des nuages de points basée sur les statistiques de co-occurrence 3D. L'approche proposée implique une stratégie de voxelisation, où le concept de matrice de co-occurrence est étendu en 3D pour calculer l'occurrence dans les 26 directions possibles. Les attributs de Haralick sont ensuite calculées en fonction de l'espace de couleur sélectionné. Une étape de régression sert à mapper ces attributs à la vérité terrain (scores subjectifs) associés aux modèles de nuages de points. Les résultats expérimentaux montrent l'efficacité de l'utilisation des statistiques de cooccurrence 3D pour l'évaluation de la qualité des nuages de points (CO-PCQA). La métrique proposée présente de bonnes performance en comparaison avec la littérature.

Mots clefs

Nuages de points, Métrique de qualité sans référence, Matrice de co-occurences.

1 Introduction

Un nuage de points 3D (NP) est une collection de points dans l'espace 3D [1], chacun avec ses coordonnées géométriques $(X, Y, Z) \in \mathbf{Z}$ et des attributs associés tels que la couleur, la réflectance et la luminance. Les nuages de points sont de plus en plus populaires ces dernières années comme moyen de représenter, stocker et compresser du contenu 3D pour diverses applications telles que la navigation, l'industrie et la réalité virtuelle. Cependant, en raison du grand nombre de points et du potentiel de distorsions lors de l'acquisition, du stockage ou de la compression, il est important de mesurer l'impact sur la qualité visuelle. L'évaluation de cette dernière par des métriques, idéalement sans référence, semble être une alternative pratique. Pour surmonter ce défi, il est nécessaire d'extraire directement des attributs qui soient corrélées à la perception humaine à partir de nuages de points dégradés. La tâche est loin d'être aisée à cause de la nature même du nuage de points, constitué de points désordonnés dans l'espace 3D. En tant que domaine émergent, différentes métriques avec référence ont été proposées pour les nuages de points. Deux métriques largement utilisées sont P2point et P2plane [2], qui comparent un nuage de points dégradé à un nuage de points de référence en mesurant leur distance géométrique. Cependant, ces métriques sont purement géométriques et ne tiennent pas compte des caractéristiques de couleur ou de texture, qui jouent souvent un rôle significatif dans la perception de la qualité. Pour pallier cette limitation, Gabriel et al. [3] ont proposé la métrique PC-MSDM, inspirée de SSIM [4], qui extrait et compare les caractéristiques visuelles du nuage de points dégradé et de sa référence. Cette métrique ne considère que des caractéristiques structurelles mais sa version étendue, PCQM [5], inclut des caractéristiques de couleur. Une autre approche consiste à utiliser les motifs binaires locaux (LBP de l'anglais Local Binary Patterns) [6], pour mesurer les différences de motifs texturaux et les traduire en terme de qualité visuelle [7]. D'autres chercheurs ont proposé de projeter les nuages de points 3D sur différents plans, comme c'est souvent le cas lors de tests subjectifs, pour obtenir des images 2D afin de les traiter avec des métriques 2D [8].

Plusieurs métriques sans référence ont été proposées pour évaluer la qualité des nuages de points. Une des approches consiste à extraire des attributs basés sur la géométrie et la couleur pour entraîner des modèles d'apprentissage automatique [9]. D'autres approches exploitent les réseaux de neurones convolutionnels et les blocs résiduels en tant qu'extracteurs d'attributs et appliquent la technique de projection multi-vue pour obtenir des projections 2D sous différents angles avant d'entraîner le modèle CNN [10]. Ces méthodes sont encore limitées par la dépendance aux méthodes de rendu, la perte d'informations due la projection et les différences entre les contenus 3D et 2D.

Dans cet article, nous proposons une nouvelle approche basée sur l'analyse de motifs de texture en utilisant une version étendue de la matrice de co-occurrences. Cela implique une voxelisation du nuage de points pour obtenir une représentation 3D régulière et le calcul de la matrice de cooccurrence 3D pour extraire les attributs de Haralick, qui sont utilisées pour entraîner des modèles d'apprentissage automatique pour prédire la qualité du nuage de points.

2 Approche proposée

La métrique proposée, baptisée CO-PCQA, suit le schéma donné par la figure 1 et est décrite ci-dessous :



FIGURE 1 – Diagramme de la métrique de qualité des nuages de points basée sur la matrice de co-occurrences (CO-PCQA). N représente le nombre de niveaux de gris.

2.1 Conversion couleur

Dans cette étude, nous avons considéré 3 espaces colorimétriques différents effectuées avant l'étape de voxelisation. Tout d'abord, l'espace RVB utilisé dans diverses applications. Ensuite l'espace YUV aui permet de décorréler la luminance et la chrominance. Plusieurs modèles négligent la chrominance et se concentrent uniquement sur la luminance car elle a un impact plus important sur le Système Visuel Humain (SVH) [11]. Enfin, l'espace colorimétrique gaussien [12], considéré comme plus représentatif du fonctionnement du SVH. Ce dernier est formalisé comme suit :

$$\begin{bmatrix} \hat{E} \\ \hat{E}_{\lambda} \\ \hat{E}_{\lambda\lambda} \end{bmatrix} = \begin{pmatrix} 0.06 & 0.63 & 0.27 \\ 0.30 & 0.04 & -0.35 \\ 0.34 & -0.6 & 0.17 \end{pmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$
(1)

où \hat{E} représente la luminance et \hat{E}_{λ} , $\hat{E}_{\lambda\lambda}$ deux canaux de chrominance.

2.2 Voxelisation



FIGURE 2 – Illustration du processus de voxelisation, pour convertir la représentation d'objet 3D à partir de points distribués dans l'espace en une matrice régulière 3D (les cubes blancs représentent les trous).

L'objectif principal de la voxelisation est de convertir un nuage de points, qui est un ensemble irrégulier de points, en un domaine cubique continu. Nous commençons par sélectionner la résolution (taille du voxel) pour déterminer la taille de la matrice résultante en fonction de la boîte englobante du nuage de points et de la résolution souhaitée. Le résultat de ce processus est une matrice tridimensionnelle, ainsi qu'un masque 3D qui indique si un voxel est vide, comme représenté dans la Figure 2.

2.3 Co-occurrence matrix

La matrice de co-occurrence est une matrice de dépendance spatiale des niveaux de gris qui caractérise la texture de l'image en quantifiant la fréquence d'apparition de paires de pixels ayant une valeur spécifique et une relation spatiale spécifiée, telle que la distance et l'angle [13]. C'est un extracteur de caractéristiques utile pour les problèmes basés sur la texture, y compris l'évaluation de la qualité.

Pour les images 2D, la matrice de co-occurrence est calculée comme suit : Pour une image en niveaux de gris I, la matrice de co-occurrence est une matrice carrée (C)de taille N, où N est le nombre de niveaux de gris dans l'image. La $(i, j)^{me}$ cellule de C représente le nombre de fois où un pixel P_1 avec une valeur d'intensité L_{P_1} est connecté à un pixel P_2 avec une valeur d'intensité L_{P_2} à une distance particulière k dans la direction θ .

Un pixel P peut avoir jusqu'à 8 pixels environnants, représentant 8 directions différentes dans une image 2D. Cependant, dans les nuages de points 3D, les points environnants ne correspondent pas nécessairement à des angles ou des directions spécifiques, mais plutôt à des points voisins qui se trouvent dans un certain rayon ou seuil de distance. Par conséquent, le nombre de directions possibles dans les nuages de points 3D dépend de la géométrie de la distribution de points et du seuil de distance choisi.

La voxelisation aide à pallier ce problème complexe en permettant de calculer la matrice de co-occurrence dans les 26 directions possibles en 3DCela donne lieu à 26 matrices de taille $(N \times N)$, où N représente le nombre de niveaux de gris. Chaque matrice 2D extraite de la matrice 3D correspond à l'une des 26 directions et est utilisée pour extraire les attributs de Haralick.

2.4 Attributs de Haralick et qualité

Les attributs de Haralick, telles que l'entropie, le contraste et le moment angulaire (ASM) de deuxième ordre, sont connus pour leur capacité à décrire les textures et leur corrélation avec la qualité visuelle. Cependant, dans ce cas particulier, les attributs sont calculés séparément pour chaque direction afin de permettre une plus grande précision, car chacune des 26 directions peut avoir un impact distinct sur la perception visuelle et potentiellement sur la qualité visuelle. Les attributs sont calculés à partir de la matrice de co-occurrence de niveaux de gris de chaque direction, ce qui améliore encore la précision de l'analyse.

Dans notre étude, nous avons considéré 16 attributs pour chaque direction, ce qui donne un total de 416 pour chaque NP. Ces attributs sont utilisés comme entrée d'un régresseur pour les mapper à la qualité finale en utilisant des scores subjectifs.

La métrique de qualité proposée utilise les attributs précédemment extraits ainsi qu'une étape de régression. Nous avons adopté trois méthodes de régression différentes, à savoir la régression par forêt d'arbres décisionnels [14], la régression par gradient boosting [15] et la régression par vecteur de support [16]. De plus, nous avons mis en un perceptron multicouches, qui prend le vecteur d'attributs en entrée et estime le score de qualité final à l'aide des scores subjectifs fournis par les différents ensembles de données.

3 Expérimentations

3.1 Configuration expérimentale

Bases de données. Pour évaluer l'efficacité de CO-PCQA, nous avons utilisé deux bases de données largement reconnues comprenant des scores de qualité subjective. Ces bases de données sont décrites ci-dessous :

- SJTU [17] est basée sur 9 nuages de points MPEG dégradés avec 7 types de distorsions (bruit Gaussien, compression, etc.) à 6 niveaux différents.
- WPC contient 20 objets et 700 versions de chaque objet altérées par différentes distorsions (souséchantillonnage, bruit Gaussien, compression, etc.)

Évaluation de la performance. Pour évaluer la performance de la prédiction, le score prédit est comparé à la vérité terrain en utilisant le coefficient de corrélation de Pearson (PLCC) pour la précision, et le coefficient de corrélation de Spearman (SRCC) pour la monotonie. La performance des métriques est évaluée en utilisant le principe de Pareto, avec 80% des données pour l'entraînement et les 20% restants pour le test. Cette procédure est répétée 1000 fois, et le résultat médian des itérations est considéré comme le résultat final et rapporté dans le Tableau 1.

3.2 Résultats

Pour analyser la performance de notre métrique proposée CO-PCQA, nous fournissons des scores notés CO-PCQA-* où -SVR, -RF, -GB et -MLP représentent l'utilisation de la régression par vecteurs de support, de la forêt d'arbres décisionnels, du gradient boosting et du perceptron multicouches, respectivement. Nous comparons nos scores à ceux obtenus à l'aide d'autres métriques de qualité avec référence et sans référence, comme indiqué dans le tableau 1. La comparaison est effectuée en utilisant l'espace de couleur YUV et une taille de voxel de 2.

Globalement, CO-PCQA donne les meilleurs résultats sur les bases SJTU et WPC avec un écart significatif entre le niveau de corrélation. Cette différence pourrait s'expliquer par le fait que la complexité et la diversité des nuages de points sont plus élevées dans WPC que dans SJTU.

TABLEAU 1 – Performance de CO-PCQA comparée aux métriques FR et NR de l'état de l'art sur les bases de données SJTU-PCQA [17] et WPC [1]. Les meilleurs résultats sont en gras et les seconds meilleurs sont soulignés.

Dataset /		SJTU-PCQA		WPC	
Approach	Metric	PLCC	SRCC	PLCC	SRCC
Avec référence	P2point	0.65	0.62	0.43	0.41
	P2plane	0.66	0.59	0.40	0.37
	GraphSIM	0.59	0.57	0.75	0.75
	PCQM	0.86	0.84	0.44	0.44
	LP-PCQM	<u>0.90</u>	0.88	0.71	0.72
Sans réference	ResSCNN	0.58	0.56	-	-
	IT-PCQA	-	-	0.55	0.54
	PQA-Net	0.85	0.82	0.70	0.69
	NR-PCQA	0.88	0.87	0.65	0.64
	CO-PCQA-SVR	0.82	0.78	0.69	0.75
	CO-PCQA-RF	0.89	0.79	0.71	0.74
	CO-PCQA-GB	0.90	0.90	0.79	0.79
	CO-PCQA-MLP	0.91	<u>0.89</u>	<u>0.78</u>	<u>0.76</u>

En se basant sur les métriques de performance présentées dans le tableau, la méthode CO-PCQA démontre une supériorité claire par rapport aux métriques sans référence testées, et fonctionne de manière comparable ou meilleure que les métriques avec référence. Les meilleures performances sont obtenues avec la méthode basée sur MLP sur la base de données SJTU, avec des valeurs PLCC et SRCC de 0,91 et 0,89, respectivement. Sur la base de données WPC, le gradient boosting obtient les meilleurs résultats avec des valeurs PLCC et SRCC de 0,79 chacune. Ces résultats indiquent que l'utilisation des statistiques de cooccurrences et des attributs de Haralick peut prédire efficacement la qualité des nuages de points. Parmi les métriques de pointe, NR-PCQA et PQA-Net fournissent de bonnes performances sur les deux bases de données, mais restent globalement en dessous de CO-PCQA-GB et CO-PCQA-MLP. Les méthodes basées sur des points tels que P2point et P2plane de MPEG sont moins efficaces car elles ne considèrent que la distribution de points et ne tiennent pas compte des distorsions de couleur ou de l'information structurelle. Enfin, ResSCNN et IT-PCQA montrent les pires performances parmi les métriques sans référence. Comparée aux méthodes précédemment mentionnées, CO-PCQA offre plusieurs avantages. Tout d'abord, elle utilise efficacement les caractéristiques de Haralick pour décrire la structure texturale d'un objet, ce qui améliore la précision du modèle. De plus, CO-PCQA considère les 26 directions individuellement, plutôt que de se fier à des caractéristiques globales, ce qui donne un modèle plus solide avec une précision améliorée. Dans l'ensemble, ces caractéristiques font de CO-PCQA une méthode très efficace pour l'analyse texturale des nuages de points.

3.3 Ablation study

Pour mieux comprendre le comportement de la métrique CO-PCQA proposée, nous avons mené une étude par abla-

tion afin d'examiner l'impact des différents paramètres impliqués dans la métrique. A cause de la limitation de page, nous ne donnerons que les résultats les plus importants utilisant la version basée sur SVR.

Taille du voxel. C'est un paramètre critique dans CO-PCQA car il détermine la précision du format voxelisé. Des voxels de taille 2, 3, 5, et 10 ont été testé et performance inversement proportionnelle a été notée. Ce qui était prévisible au vu de la variation de la précision du nuage de points. Un compromis entre performance et complexité devra être trouvé en fonction de l'application ciblée.

Représentation de la couleur. L'espace couleur peut avoir un impact sur la performance des métriques de qualité. Ici, nous évaluons la performance de l'utilisation d'une luminance seule et de trois espaces couleur différents à savoir RVB, YUV et GCS. Les résultats ont montré une similarité des performances avec un léger avantage pour GCS et une nette supériorité par rapport à la luminance seule.

Direction des co-occurrences. Cela représente l'angle d'occurrence entre un voxel et ses voisins. Nous avons ainsi combiné les angles compatibles et mesuré les performances associées. Il s'avère que l'utilisation de certaines directions privilégiées de manière indépendante génère une baisse significative de performance par rapport aux résultats du Tableau 1. Cela indique que chaque direction apporte une force supplémentaire à l'approche globale.

4 Conclusion

Cet article présente une nouvelle métrique appelée CO-PCQA qui évalue la qualité des nuages de points sans référence, basée sur des matrices de co-occurrence. Nous avons adapté ces matrices 2D basées aux nuages de points 3D et calculé les attributs de Haralick en fonction des angles d'occurrence. Nous avons expérimenté plusieurs méthodes de régression, notamment SVR, Random Forest, Gradient Boosting et Multilayer Perceptron. Les résultats expérimentaux montrent que notre modèle surpasse la plupart des métriques de référence complète et celles sans référence. Les meilleurs résultats ont été obtenus avec GB et MLP pour SJTU et WPC, respectivement. Bien que le coût de calcul de cette méthode soit élevé en raison des matrices multidimensionnelles et des caractéristiques statistiques à calculer, en plus de l'étape de voxelisation, les résultats obtenus montrent que la métrique fournit des prédictions précises sans aucune référence. Un compromis doit être trouvé entre le coût de calcul et la précision de la performance, en fonction de l'application et/ou de l'appareil ciblé. Plusieurs aspects restent à explorer, tels que l'incorporation de la saillance lors du calcul des matrices de co-occurrence.

Références

 H.L. Su, Z. F. Duanmu, W. Liu, Q. Liu, et Z. Wang. Perceptual quality assessment of "3d" point clouds. Dans *IEEE ICIP*, pages 3182–3186, 2019.

- [2] R.N. Mekuria, Z. Li, C. Tulvan, et P. Chou. Evaluation criteria for pcc (point cloud compression). 2016.
- [3] G. Meynet, J. Digne, et G. Lavoué. Pc-msdm : A quality metric for 3d point clouds. Dans *IEEE QoMEX*, pages 1–3, 2019.
- [4] A. Hore et D. Ziou. Image quality metrics : Psnr vs. ssim. Dans *IEEE ICPR*, pages 2366–2369, 2010.
- [5] G. Meynet, Y. Nehmé, J. Digne, et G. Lavoué. Pcqm : A full-reference quality metric for colored 3d point clouds. Dans *IEEE QoMEX*, pages 1–6, 2020.
- [6] R. Diniz, P. G. Freitas, et M. Farias. Towards a point cloud quality assessment model using local binary patterns. Dans *IEEE QoMEX*, pages 1–6, 2020.
- [7] P. G. Freitas, S. Alamgeer, et M. Akamine, W.and Farias. Blind image quality assessment based on multiscale salient local binary patterns. Dans ACM Multimedia Systems, pages 52–63, 2018.
- [8] Q. Yang, H. Chen, Z. Ma, Y. Xu, R. Tang, et J. Sun. Predicting the perceptual quality of point cloud : A 3d-to-2d projection-based exploration. *IEEE Trans.* on Multimedia, 2020.
- [9] Z. Zhang, W. Sun, X. Min, T. Wang, W. Lu, et G. Zhai. No-reference quality assessment for 3d colored point cloud and mesh models. *arXiv preprint arXiv* :2107.02041, 2021.
- [10] Qi Liu, Hui Yuan, Honglei Su, Hao Liu, Yu Wang, Huan Yang, et Junhui Hou. Pqa-net : Deep no reference point cloud quality assessment via multi-view projection. *IEEE TCSVT*, 31(12) :4645–4660, 2021.
- [11] M. Khosravy, N. Gupta, N. Marina, I. K Sethi, et M.R. Asharif. Perceptual adaptation of image based on chevreul-mach bands visual phenomenon. *IEEE Signal Processing Letters*, 24(5):594–598, 2017.
- [12] J.-M. Geusebroek, R. Van Den Boomgaard, A. WM Smeulders, et A. Dev. Color and scale : The spatial structure of color images. Dans *ECCV*, pages 331– 341, 2000.
- [13] L. Nanni, S. Brahnam, S. Ghidoni, E. Menegatti, et T. Barrier. Different approaches for extracting information from the co-occurrence matrix. *PloS one*, 8(12), 2013.
- [14] L. Breiman. Random forests. *Machine learning*, 45:5–32, 2001.
- [15] T. Chen et C. Guestrin. Xgboost : A scalable tree boosting system. Dans ACM int. Conf. on knowledge discovery and data mining, pages 785–794, 2016.
- [16] O.L. Mangasarian et D.R. Musicant. Robust linear and support vector regression. *IEEE TPAMI*, 22(9):950–955, 2000.
- [17] Y. Liu, Q. Yang, Y. Xu, et L. Yang. Point cloud quality assessment : Large-scale dataset construction and learning-based no-reference approach. *arXiv preprint arXiv :2012.11895*, 2020.