

Sobre o Impacto da Compressão de Nuvens de Pontos LiDAR na Segmentação Semântica Remota

Tiago de S. Fernandes e Ricardo L. de Queiroz

Resumo— Os veículos autônomos dependem de sensores LiDAR para gerar nuvens de pontos 3D, garantindo segmentação precisa e detecção de objetos. No contexto de uma cidade inteligente, queremos entender o efeito que a transmissão (compressão) pode ter na segmentação remota (em nuvem), em vez do processamento local. Neste breve artigo, buscamos compreender o impacto da compressão de nuvens de pontos no desempenho da segmentação semântica e estimar os requisitos necessários de largura de banda. Desenvolvemos uma nova métrica de distorção para avaliar esse impacto. Foram testados dois algoritmos de compressão do MPEG (GPCC e L3C2) e dois algoritmos líderes de segmentação semântica (2DPASS e PVKD) no conjunto de dados Semantic KITTI. Os resultados indicam que uma alta qualidade de segmentação exige uma taxa de comunicação de aproximadamente 0,6 MB/s para G-PCC e 2,8 MB/s para L3C2. Esses resultados são importantes para o planejamento de recursos de infraestrutura na navegação autônoma.

Abstract— Autonomous vehicles rely on LiDAR sensors to generate 3D point clouds for accurate segmentation and object detection. In a context of a smart city framework, we would like to understand the effect that transmission (compression) can have on remote (cloud) segmentation, instead of local processing. In this short paper, we try to understand the impact of point cloud compression on semantic segmentation performance and to estimate the necessary bandwidth requirements. We developed a new (suitable) distortion metric to evaluate such an impact. Two of MPEG's compression algorithms (GPCC and L3C2) and two leading semantic segmentation algorithms (2DPASS and PVKD) were tested over the Semantic KITTI dataset. Results indicate that high segmentation quality requires communication throughput of approximately 0.6 MB/s for G-PCC and 2.8 MB/s for L3C2. These results are important in order to plan infrastructure resources for autonomous navigation.

Palavras-Chave— nuvens de pontos lidar, lidar point cloud, compressão, compression, segmentação semântica, semantic segmentation.

I. INTRODUCTION

Os veículos autônomos dependem de sensores de detecção e alcance de luz (LiDAR) para perceber seu entorno [1]. Esses sensores geram enormes quantidades de dados para criar nuvens de pontos 3D (PC) detalhadas, que são submetidas à segmentação semântica e à detecção de objetos. Para processar, transmitir ou armazenar esses dados, a compressão é essencial [2]. Organizações internacionais, como o Moving Picture Experts Group¹ (MPEG), estão ativamente desenvolvendo padrões de compressores de nuvens de pontos [3].

T. Fernandes estava no Departamento de Engenharia Elétrica da Universidade de Brasília, Brasília, Brasil, e-mail: tiagotsf2000@gmail.com ; R. de Queiroz está no Departamento de Ciências da Computação da Universidade de Brasília, Brasília, Brasil, e-mail: queiroz@ieec.org.

¹www.mpeg.org

Uma etapa importante no processamento de sistemas autônomos é a segmentação semântica, uma técnica essencial de visão computacional para interpretar nuvens de pontos. Ela atribui uma classe semântica a cada ponto (por exemplo, carro, pessoa, estrada), permitindo que os sistemas compreendam seu ambiente e tomem decisões críticas em tempo real. A integridade e qualidade dos dados são fundamentais para garantir a precisão dos algoritmos de segmentação semântica [4], o que pode impactar na segurança do sistema de condução autônoma. Nosso objetivo é avaliar o impacto da compressão de nuvens de pontos no desempenho dos algoritmos de segmentação semântica. Um dos desafios centrais dessa avaliação é a seleção adequada de uma métrica de distorção. Idealmente, o objetivo seria reduzir ou eliminar fatalidades, mas não há algoritmos consolidados que relacionem a qualidade da segmentação aos padrões de condução autônoma. Portanto, buscamos preservar a qualidade do algoritmo de segmentação, ou seja, minimizar as modificações nas classes segmentadas causadas pelos algoritmos de compressão. As métricas convencionais de distorção para nuvens de pontos pressupõem correspondência entre pontos e não são adequadas por vários motivos. Por exemplo, pontos podem desaparecer ou mudar de classe, e objetos podem desaparecer, aparecer ou alterar sua forma. Também queremos levar em consideração se há humanos envolvidos na mudança de classe.

Utilizamos o popular conjunto de dados Semantic KITTI [5] para os testes. Seu “site” hospeda uma competição de segmentação semântica², da qual selecionamos dois dos melhores desempenhos para testes: O esquema de segmentação agregada baseado em projeção 2D ou 2D *Projection-based Aggregated Segmentation Scheme* (2DPASS) [6] e a destilação de conhecimento ponto-vóxel ou *Point-Voxel Knowledge Distillation* (PVKD) [7]. Assumimos que esses métodos representam o estado da arte em segmentação para nossos propósitos e incentivamos o leitor a consultar as referências para mais detalhes. Para comprimir as nuvens de pontos, utilizamos dois algoritmos de compressão do MPEG: *Geometry-based Point Cloud Compression* (G-PCC) [8] e o *Low-Latency Low-Complexity Codec* (L3C2) [9]. G-PCC é mais sofisticado e empregamos codificação octree para a geometria e RAHT [10] para os atributos. L3C2, por outro lado, é um algoritmo mais simples, desenvolvido para atender às necessidades específicas das nuvens de pontos LiDAR em aplicações em tempo real e com restrições de energia. Nosso objetivo é encontrar o nível máximo de compressão que pode ser aplicado por

²<http://www.semantic-kitti.org/tasks.html#semseg>. O acesso foi feito em outubro de 2023.

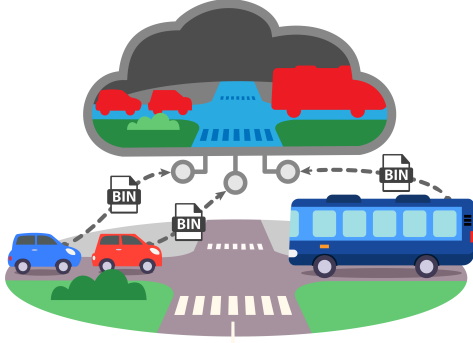


Fig. 1

CENÁRIO PARA SEGMENTAÇÃO SEMÂNTICA REMOTA DE PCs LiDAR.

cada codificador antes que a qualidade da segmentação seja substancialmente degradada.

O impacto da compressão na segmentação de nuvens de pontos também foi investigado recentemente [11], [12], [13]. No entanto, consideramos compressores e segmentadores mais atualizados e propomos uma nova métrica de distorção. Os resultados ainda são incipientes para aplicações práticas, e acreditamos que todas essas análises sejam complementares.

II. PROPOSTA DE MÉTRICA DE AVALIAÇÃO DE SEGMENTAÇÃO SEMÂNTICA

As métricas D1 e D2 [14] são comumente utilizadas para avaliar a compressão e a precisão da reconstrução de nuvens de pontos. A métrica D1 (erro ponto a ponto) mede a distância média entre os pontos correspondentes na nuvem de pontos original e na reconstruída. A métrica D2 (erro ponto a plano) mede a distância média de cada ponto na nuvem reconstruída até o plano definido pelo seu ponto correspondente e seus vizinhos na nuvem original.

Avaliar a qualidade da segmentação comparando-a com o *ground truth* não é trivial. Isso ocorre porque a nuvem de pontos comprimida pode estar distorcida e conter menos pontos do que a original, o que leva à perda da correspondência direta de rótulos.

Nossa abordagem oferece um método confiável para comparar rótulos apesar dessas distorções e reduções de pontos. Para cada ponto na nuvem de pontos não comprimida, identificamos o ponto mais próximo na nuvem comprimida e comparamos seus rótulos. A métrica de erro baseia-se na proporção de discrepâncias para um determinado rótulo. Importante ressaltar que essa métrica é computada apenas na direção da nuvem de pontos não comprimida para a nuvem de pontos comprimida, de modo a evitar que o número de pontos na nuvem comprimida influencie diretamente o resultado. Uma segmentação ainda pode representar perfeitamente todas as informações presentes na nuvem de pontos original, mesmo que haja menos pontos disponíveis.

Dada a nuvem de pontos não comprimida X , a nuvem comprimida Y e um rótulo λ , seja X_i um ponto em X e Y_{X_i} o ponto mais próximo de X_i em Y , usando a distância euclidiana. Seja $L(\cdot)$ a função que retorna o rótulo do ponto

fornecido. Defina o conjunto S_λ como o conjunto contendo todos os pares correspondentes envolvendo o rótulo λ :

$$S_\lambda = \{(X_i, Y_{X_i}) \mid (L(X_i) = \lambda) \text{ or } (L(Y_{X_i}) = \lambda)\}. \quad (1)$$

Nós similarmente definimos o subconjunto de S_λ onde há desencontro de rótulos, ou seja:

$$E_\lambda = \{S_\lambda \mid L(X_i) \neq L(Y_{X_i})\}. \quad (2)$$

Desta forma, nossa métrica é:

$$\delta(X, Y, \lambda) = \frac{|E_\lambda|}{|S_\lambda|}. \quad (3)$$

Essa métrica é bem definida mesmo quando um rótulo inteiro é removido e não confunde distorção geométrica com erros de segmentação, pois não utiliza a distância entre pontos correspondentes. O erro δ se encaixa no intervalo $[0, 1]$, onde $\delta = 0$ significa uma segmentação perfeita (comparativa). Refinamos ainda mais a métrica, potencialmente atribuindo um peso maior a tipos específicos de erros. Por exemplo, acreditamos que transformar o rótulo humano (λ_h) em outro rótulo pode ser mais crítico do que outros erros. Para isso, introduzimos um termo α , que adiciona um peso à métrica quando um rótulo humano é convertido em um não humano. Se

$$E_h = \{S_{\lambda_h} \mid L(X_i) = \lambda_h; L(Y_{X_i}) \neq \lambda_h\} \quad (4)$$

for o conjunto de pares casados onde pontos humanos são mal rotulados após sofrerem compressão, então a métrica para rótulos humanos passa a ser:

$$\delta(X, Y, \lambda_h) = \frac{|E_{\lambda_h}| + \alpha \cdot |E_h|}{|S_{\lambda_h}| + \alpha \cdot |E_h|}. \quad (5)$$

Note que para $\alpha = 0$ a métrica é a mesma dos outros rótulos. Para outros rótulos, sem ser o de humanos (λ_h), a métrica é a da Eq. (3).

III. RESULTADOS

No *dataset* Semantic KITTI, cada *scan* de uma nuvem de pontos vem com as 3 coordenadas de geometria, um atributo (reflectância), e um rótulo de segmentação. São 28 classes de rótulos incluindo *road*, *sidewalk*, *terrain*, *vegetation*, *human*, etc. Seis sequências (11, 12, 13, 14, 15, e 18) foram arbitrariamente escolhidas para testes, usando 101 *scans* de cada. Usamos segmentadores 2DPASS e PVKD com seus parâmetros *default*.

As *common test conditions* (CTC) do MPEG [15] prevêm configurações de compressão para cada compressor, rotulados de *R01* a *R06*, onde *R06* leva à melhor qualidade e acarreta menos compressão. As configurações definem um parâmetro de quantização (QP) para compactar os atributos transformados, juntamente com algumas configurações gerais. Além disso, as configurações afetam a geometria ao quantizar diretamente as posições geométricas com uma Escala de Quantização (QS). Essencialmente, a geometria é multiplicada por QS e

TABELA I
Throughput MÉDIO E Quantization Scale PARA CADA TAXA DE
COMPRESSÃO.

Compressão	Throughput (MB/s)		
	G-PCC	L3C2	QS
R01	0.0408	0.1661	0.0019
R02	0.1102	0.4125	0.0039
R03	0.5729	1.7662	0.0160
R04	1.0965	2.7951	0.0310
R05	2.2189	4.2298	0.1300
R06	2.8484	4.8205	0.2500
Sem compressão	28.7465	28.7465	1.0000

arredondada. Em alguns casos, isso é semelhante à poda da *octree* (árvore octal).

Compactamos e segmentamos cada PC, comparando os resultados da segmentação do PC descompactado com os resultados da segmentação do PC não compactado. Assim, a distorção é nossa métrica para comparar a segmentação com e sem o PC passando por compressão com perdas. A taxa é calculada como *throughput* em MB/s, assumindo uma taxa de captura de 10 quadros por segundo. Utilizamos arbitrariamente $\alpha = 1$ em nossos experimentos, mas qualquer outro valor poderia ser usado.

Os valores de QS e as taxas de bits médias resultantes, em todo o conjunto de dados, estão mostrados na Tabela I. Os resultados de taxa-distorção (RD) para cada combinação de codificador-segmentador são apresentados nas Figuras 2, 3, 4 e 5, para alguns rótulos populares de segmentação. Esses gráficos mostram as médias ao longo do conjunto de dados. As taxas de bits associadas aos parâmetros de taxa $R0n$ são indicadas com linhas verticais tracejadas.

Nos gráficos de RD, tanto os compressores G-PCC quanto L3C2 demonstram resultados semelhantes, sendo que o último requer taxas de bits mais altas, como esperado. Os algoritmos de segmentação 2DPASS e PVKD exibem desempenho semelhante, mas o 2DPASS parece ser ligeiramente mais robusto contra artefatos de compressão. Para taxas baixas, a distorção é alta, pois os erros de segmentação (comparados aos PCs não compactados) são muito pronunciados. Exemplos de tal degradação são mostrados na Figuras 6, para G-PCC e 2DPASS. Elas apresentam visões dos resultados segmentados em nuvens de pontos não compactadas e compactadas usando $R03$ e $R02$. Observe que a queda na qualidade de $R03$ para $R02$ fez com que a maioria dos rótulos desaparecesse, aumentando consideravelmente a distorção.

A partir dos gráficos de RD, concluímos que taxas $R05$ e superiores fornecem segmentação quase tão boa quanto a original, mas talvez $R04$ seja suficiente para a maioria das aplicações. Há uma grande queda de distorção de $R02$ para $R03$, enquanto $R01$ e $R02$ provavelmente não são aceitáveis. Talvez $R03$ (0,57 MB/s para G-PCC ou 1,77 MB/s para L3C2) seja uma configuração aceitável, exceto para a combinação L3C2 e PVKD, na qual $R04$ (2,8 MB/s) pode ser um mínimo aceitável.

Embora tenhamos usado apenas o conjunto de dados Semantic KITTI, esperamos que outros conjuntos de dados, com resolução de varredura semelhante, produzam resultados

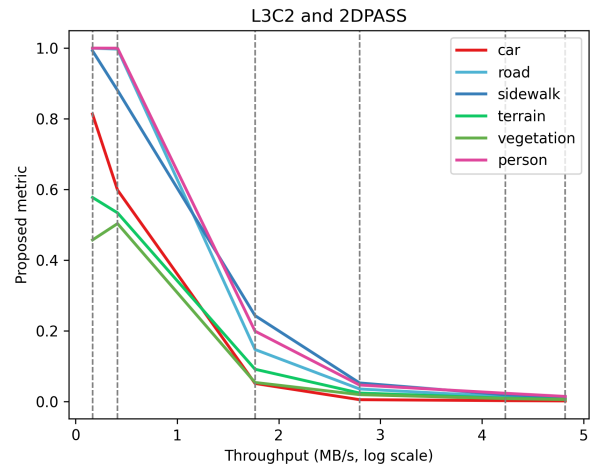


Fig. 2
RESULTADOS DA MÉTRICA PROPOSTA PARA COMPRESSÃO L3C2 E
ALGORITMO DE SEGMENTAÇÃO 2DPASS.

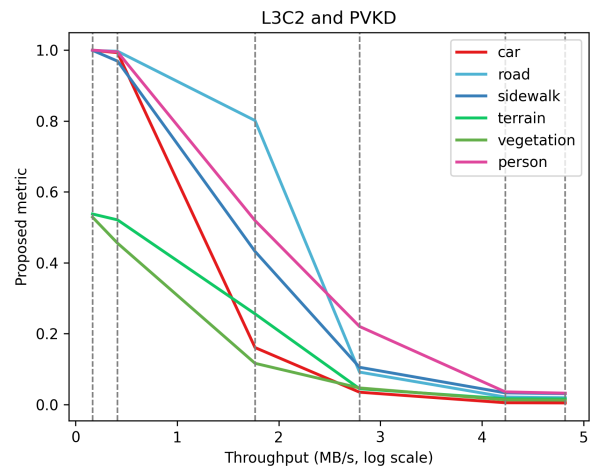


Fig. 3
RESULTADOS DA MÉTRICA PROPOSTA PARA COMPRESSÃO L3C2 E
ALGORITMO DE SEGMENTAÇÃO PVKD.

similares.

IV. CONCLUSÕES

Exploramos o impacto da compactação de nuvem de pontos nos algoritmos de segmentação semântica, com foco nos compressores G-PCC e L3C2 e nos algoritmos de segmentação 2DPASS e PVKD. Usando uma métrica nova e adequada, avaliamos como diferentes taxas de compressão afetam a qualidade da segmentação dentro do conjunto de dados Semantic KITTI. Os resultados indicam que manter uma alta qualidade de segmentação requer valores de throughput de pelo menos 0,6 MB/s para G-PCC/2DPASS e 2,8 MB/s para L3C2/PVKD.

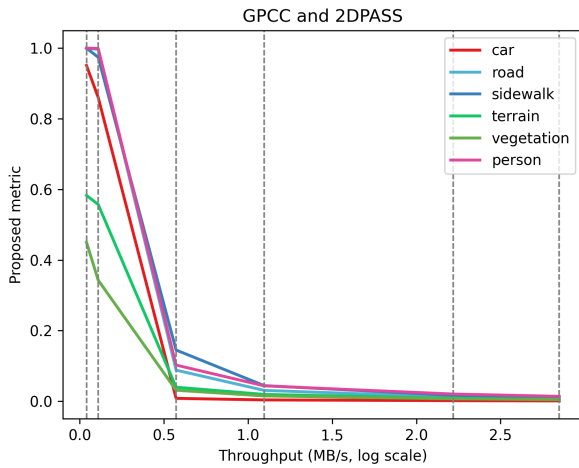


Fig. 4

RESULTADOS DA MÉTRICA PROPOSTA PARA COMPRESSÃO G-PCC E ALGORITMO DE SEGMENTAÇÃO 2DPASS.

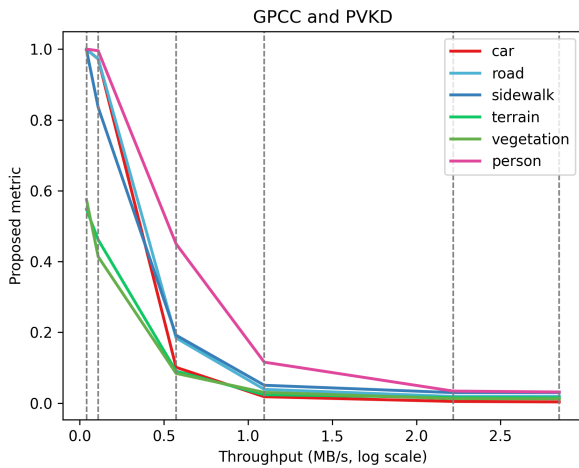


Fig. 5

RESULTADOS DA MÉTRICA PROPOSTA PARA COMPRESSÃO G-PCC E ALGORITMO DE SEGMENTAÇÃO PVKD.

Essas informações são importantes para garantir eficiência em sistemas autônomos, especialmente em cenários onde os dados são processados remotamente.

A queda significativa na qualidade da segmentação, observada na transição das configurações de compressão de $R02$ para $R04$, pode ser causada pela quantização da geometria. A redução da densidade de pontos em $R02$ elimina características importantes necessárias para uma rotulagem precisa.

Trabalhos futuros podem investigar uma compactação da geometria mais favorável à segmentação. Também precisamos incorporar alarmes de nível mais alto para quando objetos humanos aparecerem ou desaparecerem completamente.

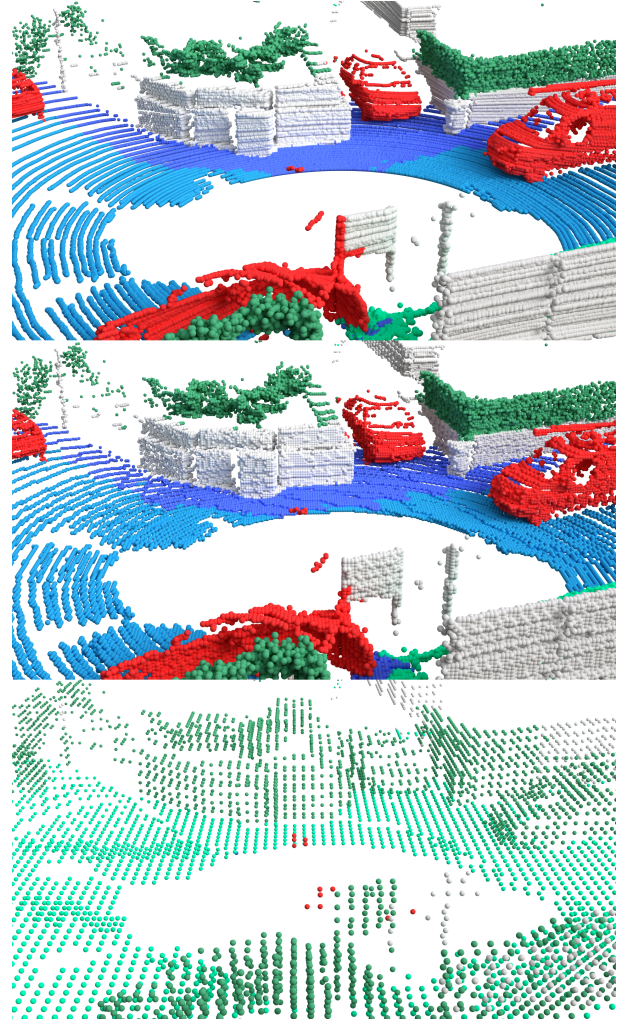


Fig. 6

QUADRO DA SEQUÊNCIA 11 USANDO G-PCC E 2DPASS. DE CIMA PARA BAIXO: SEM COMPRESSÃO, R03 E R02.

REFERÊNCIAS

- [1] Y. Li and J. Ibanez-Guzman, "LiDAR for autonomous driving: The principles, challenges, and trends for automotive LiDAR and perception systems," *IEEE Signal Processing Magazine*, vol. 37, no. 4, pp. 50–61, July 2020.
- [2] R. Abbasi, A. K. Bashir, H. J. Alyamani, F. Amin, J. Doh, and J. Chen, "LiDAR point cloud compression, processing and learning for autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 1, pp. 962–979, January 2023.
- [3] S. Schwarz, M. Preda, V. Baroncini, M. Budagavi, P. Cesar, P. A. Chou, R. A. Cohen, M. Krivokuća, S. Lasserre, Z. Li, J. Llach, K. Mammou, R. Mekuria, O. Nakagami, E. Siahaan, A. Tabatabai, A. M. Tourapis, and V. Zakharchenko, "Emerging MPEG standards for point cloud compression," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 1, pp. 133–148, March 2019.
- [4] B. Gao, Y. Pan, C. Li, S. Geng, and H. Zhao, "Are we hungry for 3D LiDAR data for semantic segmentation? a survey of datasets and methods," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6063–6081, July 2022.
- [5] J. Behley, M. Garbade, A. Milioto, J. Quenzel, S. Behnke, C. Stachniss, and J. Gall, "SemanticKITTI: A dataset for semantic scene understanding of LiDAR sequences," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, South Korea, February 2019.
- [6] X. Yan, J. Gao, C. Zheng, C. Zheng, R. Zhang, S. Cui, and Z. Li, "2DPASS: 2D priors assisted semantic segmentation on LiDAR point

- clouds,” in *European Conference on Computer Vision*, Tel Aviv, Israel, July 2022.
- [7] Y. Hou, X. Zhu, Y. Ma, C. C. Loy, and Y. Li, “Point-to-voxel knowledge distillation for LiDAR semantic segmentation,” in *IEEE Conference on Computer Vision and Pattern Recognition*, New Orleans, United States, June 2022.
 - [8] D. Graziosi, O. Nakagami, S. Kuma, A. Zaghetto, T. Suzuki, and A. Tabatabai, “An overview of ongoing point cloud compression standardization activities: video-based (V-PCC) and geometry-based (G-PCC),” *APSIPA Transactions on Signal and Information Processing*, vol. 9, p. e13, April 2020.
 - [9] S. Lasserre and J. Taquet, “A point cloud codec for LiDAR data with very low complexity and latency.” m56477. ISO/IEC JTC 1/SC 29/WG 7., 2021.
 - [10] R. L. de Queiroz and P. A. Chou, “Compression of 3D point clouds using a region-adaptive hierarchical transform,” *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3947–3956, August 2016.
 - [11] N. A. B. Martins, L. A. d. S. Cruz, and F. Lopes, “Impact of lidar point cloud compression on 3D object detection evaluated on the KITTI dataset,” *J. Image Video Process.*, vol. 2024, no. 1, June 2024.
 - [12] L. Garrote, J. Perdiz, L. A. da Silva Cruz, and U. J. Nunes, “Point cloud compression: Impact on object detection in outdoor contexts,” *Sensors*, vol. 22, no. 15, August 2022.
 - [13] M. Roque and L. A. da Silva Cruz, “An empirical study of the effect of point cloud compression on the performance of segmentation,” in *2019 8th European Workshop on Visual Information Processing (EUVIP)*, January 2019.
 - [14] D. Tian, H. Ochimizu, C. Feng, R. Cohen, and A. Vetro, “Geometric distortion metrics for point cloud compression,” in *2017 IEEE International Conference on Image Processing (ICIP)*, Beijing, China, September 2017.
 - [15] “Common test conditions for point cloud compression,” ISO/IEC JTC1/SC29 Joint WG11/WG1 (MPEG/JPEG), Gothenburg, SE, Approved WG 11 doc. N18883, July 2019.