

PROPOSTA DE ANÁLISE DO USO DO CODIFICADOR DE VÍDEO AV1 PARA COMPRESSÃO DE NUVENTES DE PONTOS DINÂMICAS

GUSTAVO REHBEIN; CRISTIANO SANTOS; GUILHERME CORRÊA; MARCELO PORTO

Universidade Federal de Pelotas

Programa de Pós-Graduação em Computação – PPGC

Video Technology Research Group - ViTech

{ghrehbein, cfdsantos, gcorrea, porto}@inf.ufpel.edu.br

1. INTRODUÇÃO

As nuvens de pontos são usadas para representar objetos e cenários em um espaço tridimensional com alta fidelidade. Elas são constituídas por uma coleção de pontos com coordenadas no espaço 3D em que associado a cada ponto pode haver informações relacionadas à aparência daquele ponto (cor, reflexividade, etc.) (Figura 1 (a)). Atualmente, nuvens de pontos são usadas em uma gama de aplicações, como modelagem 3D, monitoramento ambiental, imagens biomédicas, realidade virtual, concepção de metaverso, entre outras (CAO et al., 2019)

As nuvens de pontos dinâmicas são análogas a vídeos, sendo formadas por sequência de nuvens de pontos, de modo similar a quadros em um vídeo. Assim como vídeos digitais, a transmissão e o armazenamento de nuvens de pontos é proibitiva sem a compressão de dados. Diante disso, é primordial o desenvolvimento de técnicas de compressão para que essas nuvens possam ser armazenadas e transmitidas de forma eficiente.

O Moving Picture Expert Group (MPEG), grupo responsável por diversos padrões de compressão de vídeos amplamente usados, como o H.264/AVC (*Advanced Video Coding*), o *High Efficiency Video Coding* (HEVC) e o *Versatile Video Coding* (VVC) (BROSS et al., 2021), criou um *Call of Proposals* em 2017 com foco em compressão de nuvens de pontos. Baseado nas propostas apresentadas foi estabelecido o padrão denominado o *Video-based Point Cloud Compression* (V-PCC) para codificação de nuvens de pontos dinâmicas (GRAZIOSI, 2020).

O V-PCC permite que nuvens de pontos dinâmicas sejam comprimidas utilizando-se de métodos de compressão já estabelecidos para vídeos digitais. Para isto, inicialmente a nuvem de pontos é planificada através de um processo que divide a nuvem em diversos pedaços, chamados *patches*, permitindo a projeção para o espaço 2D (GRAZIOSI, 2020).

Para cada nuvem de pontos, é gerado um mapa de ocupação, uma imagem que indica as regiões que possuem pontos, uma imagem de geometria, que representa a profundidade de cada pixel, e uma imagem para os atributos de cor, que representam informações sobre a aparência de cada ponto (Figura 1 (b)). Além disso, informações sobre o posicionamento de cada *patch* também são geradas para permitir a reconstrução da nuvem de pontos (GRAZIOSI, 2020). Os quadros planificados, então, são comprimidos usando um codificador de vídeo. Por padrão, o V-PCC usa o codificador HEVC, mas também oferece suporte para outros codificadores como o H.264/AVC e o VVC. Como o V-PCC usa os codificadores de vídeos para codificar as nuvens planificadas, ele também possibilita o uso de outros codificadores de vídeo.

Neste trabalho será explorado o uso do codificador de vídeo AOMedia Video 1 (AV1) no V-PCC, com o objetivo de explorar novas técnicas a fim de obter uma melhor eficiência de codificação. Uma vez que o AV1 é um codificador mais atual e com eficiência de compressão maior que o HEVC (CHEN et al., 2020).

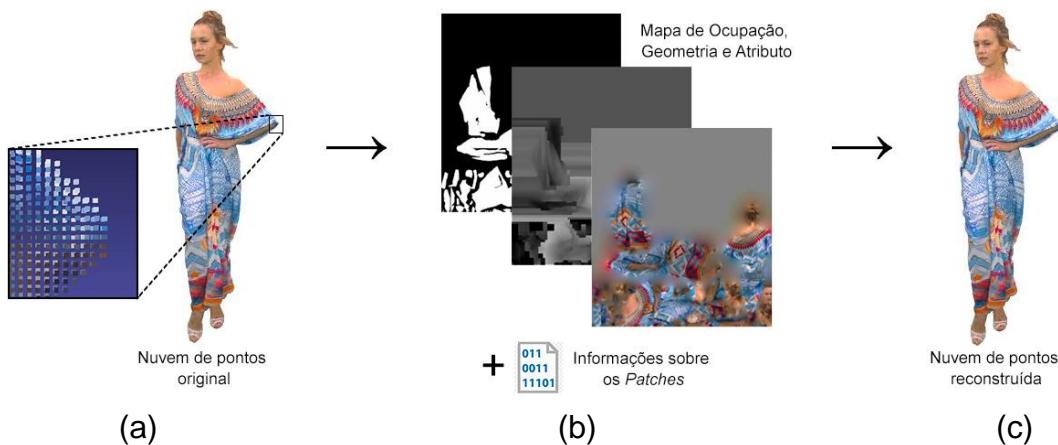


Figura 1 – Exemplo do funcionamento do V-PCC: (a) Nuvem de pontos original, (b) Processo de codificação do V-PCC e (c) Nuvem reconstruída.

2. METODOLOGIA

Para possibilitar o uso do codificador de vídeo AV1 no V-PCC, foram feitas modificações no software de referência do V-PCC (Test Model Category 2 - TMC2) (GITHUB, 2023) na versão v22.1. Para o AV1 foi usada a versão v3.6.1 do software de referência (AOMEDIA, 2023). As Condições Comuns de Teste (*Common Test Conditions* - CTC) do V-PCC (ISO/IEC, 2020) definem configurações que ajustam parâmetros tanto para o V-PCC, quanto para o codificador de vídeo usado, para cada modo de codificação (*Random Access* (RA) e *All-intra* (AI)) e cada configuração de taxa de *bits* (r1, r2, r3, r4 e r5). Assim, por se tratar de um codificador de vídeo novo, novas configurações foram criadas para uso com o AV1.

Para a realização destes experimentos foi feita uma análise sobre as configurações já especificadas no V-PCC, a fim de identificar o comportamento de cada uma. Foi observado que as configurações de *taxa de bits* controlam os parâmetros de quantização do codificador e tamanho de mapa de ocupação. Enquanto as configurações de modos alteram tanto o funcionamento do V-PCC, quanto o codificador de vídeo usado. A Tabela 1 mostra os parâmetros de quantização usados pelo HEVC (QPs) e as conversões realizadas para o AV1 (CQs) para cada configuração de *taxa de bits*. Os CQs do AV1 foram selecionados com base nos valores indicados nas CTCs do AV1 (DAEDE et al, 2020). Para os modos de codificação AI e RA, foi usada a configuração *High Latency CQP*, sugerida por DAEDE et al. (2020). Vale ressaltar que os valores foram escolhidos empiricamente e não garantem uma equivalência aos valores e configurações usadas pelo HEVC, uma vez que se trata de dois codificadores de vídeo diferentes. Vale ressaltar ainda que as CTCs do AV1 não especificam uma configuração para o modo de codificação AI. Assim, para esse modo, foi criada uma configuração onde apenas as ferramentas de predição intra quadro foram ativadas. Esses experimentos têm caráter exploratório, buscando encontrar possibilidades para trabalhos futuros.

Tabela 1 – Parâmetros de quantização usados

Config. <i>bitrate</i>	HEVC		AV1	
	QP Geometria	QP Atributo	CQ Geometria	CQ Atributo
ctc-r5	16	22	20	28
ctc-r4	20	27	24	36
ctc-r3	24	32	32	43
ctc-r2	28	37	36	47
ctc-r1	32	42	43	55

3. RESULTADOS E DISCUSSÃO

O comportamento do V-PCC com a utilização do codificador AV1 foi avaliado usando três sequências de nuvem de pontos (*longdress*, *queen* e *dancer_vox11*), definidas nas CTCs (ISO/IEC, 2020). Para isso, os primeiros 32 quadros de cada sequência foram codificados utilizando o AV1 e comparados com os resultados obtidos quando utilizado o HEVC. Para cada sequência, foram realizados experimentos considerando as cinco configurações de *taxa de bits* e os dois modos de codificação AI e RA, totalizando 30 experimentos para cada codificador de vídeo. Os experimentos foram realizados em um computador com o processador AMD Opteron 6276s de 8 núcleos e 120GB de memória RAM, cedido pelo Grupo de Arquiteturas e Circuitos Integrados (GACI) da UFPel.

A Figura 2 apresenta os resultados médios de taxa de *bits* e distorção (*Peak Signal to Noise Ratio* - PSNR) sobre as métricas ponto-a-ponto (D1) e ponto-a-plano (D2) para a informação de geometria (TIAN et al., 2017), e PSNR dos atributos de cor. Com base nos resultados é possível observar que houve ganho de eficiência na compressão da geometria e atributo no modo RA, apenas para a métrica D1 com a taxa de *bits* mais alta ouve uma pequena perda de eficiência de compressão, quando comparado com o HEVC. No modo AI, não houve ganho de eficiência de compressão em nenhum dos casos. A não equivalência entre as configurações usadas no AV1 e HEVC para esse modo podem justificar a diferença de ganhos com o modo RA.

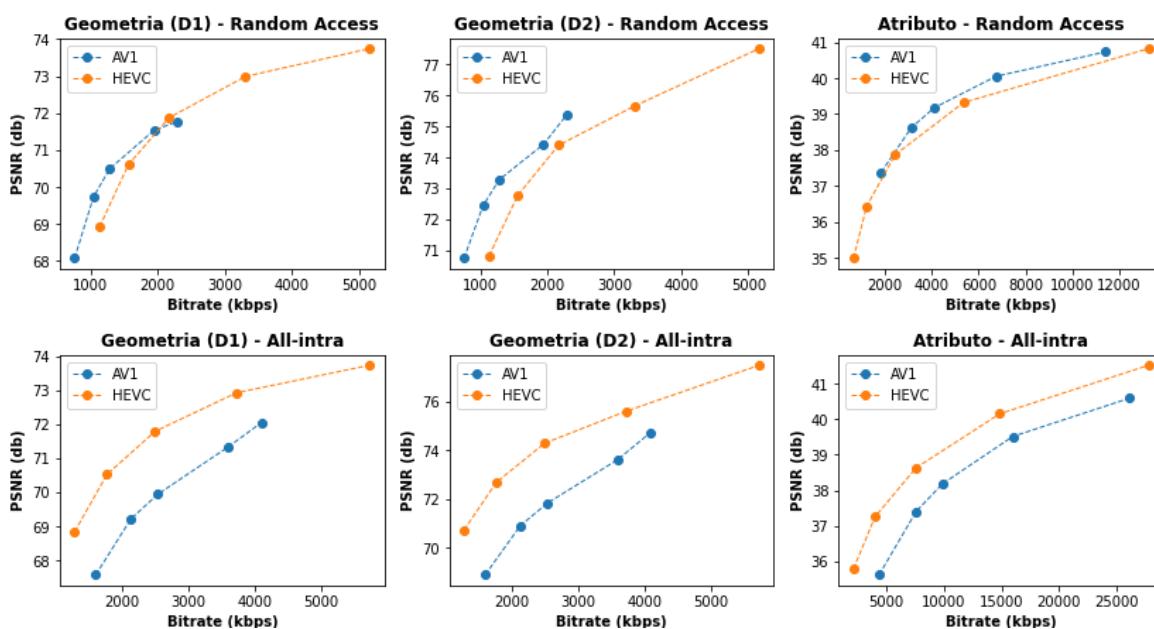


Figura 2 – Resultados médios dos experimentos.

A Tabela 2 apresenta os resultados por meio da avaliação de BD-PSNR e BD-Rate dos experimentos, onde também é observado o ganho de eficiência de compressão dos atributos e geometria no modo RA. O AV1 consegue uma compressão dos atributos de cor e geometria mais eficiente, ou seja, para uma mesma qualidade objetiva (PSNR), são necessários menos de *bits* para a representação da nuvem de pontos.

Tabela 2 - Resultados de BD-Rate e BR-PSNR médios

	Random Access			All-Intra		
	Geom. D1	Geom. D2	Atributo	Geom. D1	Geom. D2	Atributo
BD-PSNR (db)	0.41	1.01	0.06	-1.88	-2.42	-1.55
BD-RATE (%)	-14.80	-24.89	-2.54	61.21	65.85	67.33



4. CONCLUSÕES

Neste trabalho foi explorado o uso de um novo codificador de vídeo, o AV1, no codificador V-PCC, com o objetivo de explorar possibilidades de aumento de eficiência de compressão de nuvens de pontos. Modificações no software de referência do V-PCC foram realizadas para acomodar esse novo codificador. Configurações foram selecionadas para realizar uma comparação inicial entre o codificador já usado e o novo. Os experimentos realizados mostraram um significativo ganho de eficiência de compressão dos atributos de cor e geometria no modo RA.

Embora tenha sido observada uma perda de eficiência de compressão da informação de geometria e atributo no modo AI, é importante observar que, durante o desenvolvimento do V-PCC, o HEVC foi testado e otimizado para funcionar com a sua melhor eficiência possível. O fato do AV1 apresentar ganhos de eficiência com uma configuração padrão selecionada, sem testes e aprimoramentos anteriores, demonstram que, com uma implementação do AV1 focada para uso no V-PCC, os valores de eficiência de compressão podem ser ainda maiores daqueles obtidos com o HEVC. Como trabalho futuro, uma exploração das configurações do AV1 será realizada, a fim de obter uma melhor eficiência de compressão de nuvens de pontos no V-PCC.

5. REFERÊNCIAS BIBLIOGRÁFICAS

AOMEDIA. **AV1 Codec Library**. 2023. Online. Disponível em: <https://aomedia.googlesource.com/aom/>

BROSS, B. et al. Developments in international video coding standardization after avc, with an overview of versatile video coding (vvc). **Proceedings of the IEEE**, v. 109, n. 9, p. 1463-1493, 2021.

CAO, C. et al. 3D point cloud compression: A survey. In: **The 24th International Conference on 3D Web Technology**. 2019. p. 1-9.

CHEN, Y. et al. An overview of coding tools in AV1: The first video codec from the alliance for open media. **APSIPA Transactions on Signal and Information Processing**, v. 9, p. e6, 2020.

DAEDE, T. et al. Video codec testing and quality measurement. **draft-ietf-netvc-testing-09**, v. 9, p. 23, 2020.

GITHUB. **Video Point Cloud Compression - VPCC**. 2023. Online. Disponível em: <https://github.com/MPEGGroup/mpeg-pcc-tmc2>

GRAZIOSI, D. et al. 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**, v. 9, p. e13, 2020.

ISO/IEC. Coding of moving pictures and audio: Common Test Conditions for V3C and V-PCC. **ISO/IEC JTC1/SC29/WG11 Doc. N19518**, p. 10, 2020

TIAN, D. et al. Geometric distortion metrics for point cloud compression. In: **2017 IEEE International Conference on Image Processing (ICIP)**. IEEE, 2017. p. 3460-3464.