

ANÁLISE DO COMPORTAMENTO DAS FERRAMENTAS DE CODIFICAÇÃO DO VTM

OTÁVIO FERNANDES DOS SANTOS¹; ROGERIO CASSANTA ROSADO²;
LUCIANO VOLCAN AGOSTINI³

¹Universidade Federal de Pelotas – otavio.fsantos@inf.ufpel.edu.br

²Universidade Federal de Pelotas – rcassanta@inf.ufpel.edu.br

³Universidade Federal de Pelotas – agostini@inf.ufpel.edu.br

1. INTRODUÇÃO

Com o avanço tecnológico, vídeos digitais estão cada vez mais presentes na vida das pessoas e são encontrados nos mais diversos cenários com cada vez mais frequência. É estimado que até o ano 2023, os vídeos digitais irão representar 82% do tráfego total dos consumidores de internet (CISCO, 2020). Por consequência do avanço tecnológico na exibição e processamento de vídeos, com um investimento cada vez maior em vídeos de resoluções maiores e abordagens mais interativas e imersivas, vídeos omnidirecionais (360°) começam a se tornar mais frequentes e requisitados, principalmente no campo da realidade virtual (VR) (LI; XU; ZHANG; CALLET, 2015). Um dos maiores obstáculos para a popularização de vídeos omnidirecionais é uma necessidade maior de largura de banda necessária para entregar vídeos com uma qualidade aceitável.

Codificadores atualmente buscam soluções para obter ganhos na compressão de vídeos. Dentre os padrões de codificação de vídeos existentes, destaca-se o VVC (*Versatile Video Coding*), o qual foi finalizado em 2020 e é o mais recente padrão internacional de codificação de vídeos (SULLIVAN, 2020). O VVC possui um conjunto de novas funcionalidades e ferramentas, porém, a adição destas novas tecnologias, trouxe consigo um maior custo computacional, que, por consequência, necessita de novas soluções que reduzam esse custo para os diferentes processos da codificação. Dentre os vários motivos para o elevado custo computacional do VVC, cabem destacar dois de interesse deste trabalho: (i) o número elevado de tamanhos de blocos suportados para codificar os vídeos e (ii) o número de novos modos de codificação disponíveis para a predição intra-quadros (BROSS; WANG; YE; LIU; CHEN; SULLIVAN; OHM, 2021).

Com base em trabalhos realizados na área, é possível observar que vídeos omnidirecionais possuem características que tornam algumas regiões do vídeo mais susceptíveis a serem codificadas por um conjunto reduzido de ferramentas, modos e tamanhos de blocos (STORCH; CRUZ; AGOSTINI; ZATT; PALOMINO, 2019). Não é possível comprimir um vídeo omnidirecional usando diretamente um padrão como o VVC, desenvolvido para vídeos 2D convencionais. Isso ocorre porque vídeos omnidirecionais possuem uma representação esférica. Então, para codificar vídeos omnidirecionais com o codificador VVC, a representação esférica deve ser mapeada para um plano 2D. Esse processo pode ser realizado com diversos esquemas de projeção, a mais utilizada sendo a equirretangular (RAY; JUNG; LARABI, 2018). Entretanto, não é possível fazer uma projeção de uma esfera para um plano sem que distorções sejam criadas e, em uma projeção equirretangular, as distorções são mais recorrentes nas regiões polares. Com base neste comportamento, este trabalho visou explorar o comportamento das

ferramentas do codificador nas diferentes regiões de vídeos omnidirecionais, com o intuito de encontrar tendências para determinadas ferramentas de acordo com a região do vídeo analisada.

O foco da análise esteve na predição intra-quadro do VVC, que permite o uso de 17 tamanhos de bloco entre 64x64 e 4x4 amostras e diversos novos modos de predição (FRAUNHOFER, 2022). Entre os modos de predição suportados, cabe destacar: (i) predição com múltiplas linhas de referência (do inglês *multiple reference line prediction* - MRL), que permite o uso de mais de uma linha de referência; (ii) predição intra baseada em matriz (do inglês *matrix-based intra prediction* - MIP), que realiza a predição através da multiplicação de matrizes e interpolação de amostras; (iii) sub-partição intra (do inglês *intra sub-partition* - ISP), que melhora a eficiência de codificação explorando a correlação entre as amostras dentro de um bloco, subdividindo esse bloco; e (iv) para os modos normais presentes em outros codificadores: (a) Angular, que consiste em avaliar 65 opções de modos direcionais; (b) Planar, que gera a predição através de interpolação e (c) DC, realiza uma média geral das amostras do bloco (HAMIDOUCHE, 2021). A análise realizada considerou também diferentes parâmetros de quantização (do inglês *quantization parameter* – QP), conforme recomendado pelas condições comuns de teste (HANHART; BOYCE; CHOI; LIN, 2018). O QP determina a taxa de compressão e, quanto maior for o QP, maior será a compressão, mas maior também será a perda de qualidade do vídeo. Os mapas de calor foram gerados para todas as ferramentas da predição intra-quadro do VVC, para todos os tamanhos de blocos suportados e para os diferentes QPs recomendados.

2. METODOLOGIA

Inicialmente, foram feitas modificações no código software de referência do codificador VVC, chamado de *VVC Test Model* (VTM) (FRAUNHOFER, 2022), para gerar arquivos com as informações de interesse sobre a utilização das ferramentas para cada pixel dos vídeos. Então foram codificados quatro vídeos omnidirecionais no codificador VVC. Estes vídeos foram: Trolley, Chair Lift Ride, Kite Flite e Harbor. Ao utilizar mais de um vídeo, é possível obter uma melhor compreensão do comportamento do codificador. Cada vídeo foi codificado quatro vezes, usando quatro QPs: 22, 27, 32 e 37, conforme recomendado pelas condições comuns de teste (HANHART; BOYCE; CHOI; LIN, 2018). Com os resultados da codificação, foi utilizado um programa que filtra as informações e as transforma em arquivos de melhor compreensão e de mais fácil manejo. A seguir, outro programa foi usado para analisar as informações e organizar essas informações em diversos arquivos. Finalmente, um programa desenvolvido em um trabalho anterior do grupo foi usado para a geração dos mapas de calor que tornam bem simples a visualização do comportamento do codificador, de acordo com a ferramenta e a região do vídeo. Foram gerados mapas de calor por vídeo, por QP e por tamanho de bloco utilizado, mas também foram gerados mapas de calor para as médias de todos os resultados obtidos por todos os QPs de cada vídeo, para cada tamanho de bloco.

Durante a realização do trabalho foram gerados um total de 308 mapas de calor, entretanto, 108 destes mapas foram considerados mais relevantes, porque foram gerados se baseando na média entre os quatro QPs utilizados para cada vídeo. Os mapas restantes focam em vídeos e QPs específicos e, assim, podem apresentar um comportamento que não é um padrão, mas apenas um caso

isolado, e, portanto, a análise foi realizada para os 108 mapas com os resultados médios. Os mapas de calor foram gerados para cada vídeo e para cada tamanho de bloco suportado considerando oito configurações únicas: três mapas para as ferramentas ISP, MRL e MIP, um mapa para os modos normais em conjunto (Angular, Planar e DC), três mapas para os modos Angular, Planar e DC individuais e um mapa para todos os modos em conjunto.

3. RESULTADOS E DISCUSSÃO

A Figura 1 e a Figura 2 apresentam dois conjuntos de mapas de calor para o vídeo Harbor, um conjunto para tamanho de bloco 4x4 e outro para tamanho de bloco 64x64. Por uma questão de restrição de espaço, os demais mapas gerados não foram incluídos neste artigo. Nos mapas de calor apresentados na Figura 1 e na Figura 2, o azul escuro indica que a ferramenta não é usada para aquela região do vídeo e, quanto mais próximo do vermelho, maior é o percentual de uso daquela ferramenta na região respectiva do vídeo.

Na Figura 1 e na Figura 2, é possível observar que há diferença na utilização das ferramentas de acordo com o tamanho de bloco e da região do vídeo. É possível perceber que blocos menores possuem uma ocorrência focada em regiões de maiores detalhes, enquanto blocos maiores são mais frequentes em áreas homogêneas do vídeo. Também é observável que a ferramenta ISP não é utilizada com blocos pequenos, bem como a ferramenta MRL possui um comportamento irregular em blocos maiores.

Figura 1: Mapa de calor da média entre os quatro QPs para o vídeo Harbor com blocos de tamanho 4x4.

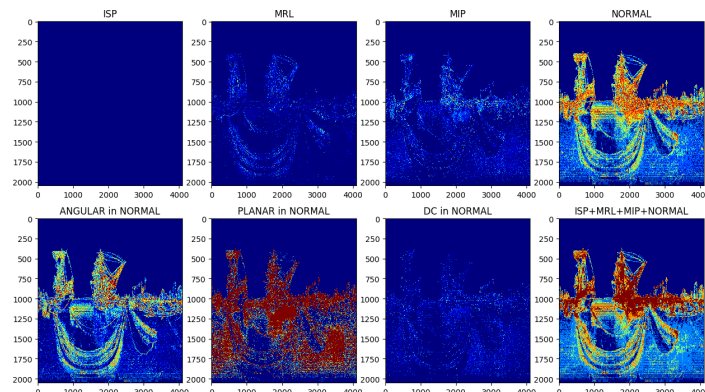
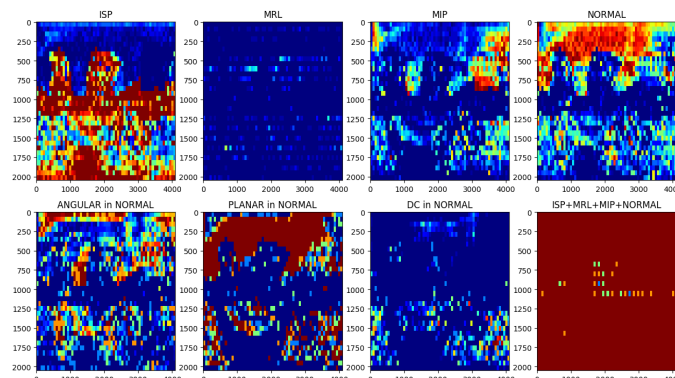


Figura 2: Mapa de calor da média entre os quatro QPs para o vídeo Harbor com blocos de tamanho 64x64.



4. CONCLUSÕES

Através dos resultados obtidos nesta investigação foi possível adquirir uma boa noção do funcionamento do codificador de acordo com as ferramentas utilizadas e tamanhos de blocos escolhidos. Com base nestes resultados, será possível propor algumas soluções para reduzir o custo computacional da predição intra-quadro do VVC quando codificando vídeos omnidirecionais.

Algumas simplificações possíveis são: (i) é possível alterar o código do codificador a fim de evitar a avaliação da ferramenta ISP em blocos menores, visto que seu uso é muito reduzido; (ii) é possível descartar o uso de blocos pequenos em regiões homogêneas do vídeo; e (iii) é possível descartar a utilização de blocos grandes em áreas heterogêneas do vídeo. Com base nestas simplificações, que são focos de futuras pesquisas, espera-se atingir bons ganhos de redução de custo computacional da predição intra-quadro do VVC quando codificando vídeos omnidirecionais.

5. REFERÊNCIAS BIBLIOGRÁFICAS

CISCO. **Cisco Annual Internet Report (2018-2023) White Paper**. Acessado em 31 julho 2022. Online. Disponível em: <https://bit.ly/3vFXV20>

LI, C; XU, M; ZHANG, S; CALLET, P. State-of-the-art in 360 Video/Image Processing: Perception, Assessment and Compression, **IEEE**, 2015.

SULLIVAN, G. Versatile Video Coding (VVC) Arrives, **IEEE**, 2020.

BROSS, B; WANG, Y; YE, Y; LIU, S; CHEN, J; SULLIVAN, G; OHM, J. Overview of the Versatile Video Coding (VVC) Standard and its Applications, **IEEE**, 2021.

STORCH, I; CRUZ, S; AGOSTINI, L; ZATT, B; PALOMINO, D. **The Impacts of Equirectangular 360-degrees Videos in the Intra-Frame Prediction of HEVC**. 2019. Journal of Integrated Circuits and Systems.

RAY, B; JUNG, J; LARABI, M. A Low-Complexity Video Encoder for Equirectangular Projected 360 Video Content. **IEEE**, 2018.

FRAUNHOFER. **VVCSoftware_VTM**. Acessado em 4 agosto 2022. Online. Disponível em: https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM

HAMIDOUCHE, W. **Versatile video coding standard**. arXiv preprint arXiv, 2021.

HANHART, P; BOYCE, J; CHOI, K; LIN, J. **JVET common test conditions and evaluation procedures for 360 videos**. 2018. JVET-L1012.