

## A CODIFICAÇÃO DE VÍDEOS 360° NO PADRÃO VVC E A FERRAMENTA INTRA SUBPARTITION (ISP) VERTICAL

JOÃO PEDRO LEAL GÓMEZ DE VARGAS; LUCIANO VOLCAN AGOSTINI

<sup>1</sup>Universidade Federal de Pelotas (UFPEL) – joao.plgdv@inf.ufpel.edu.br

<sup>2</sup>Universidade Federal de Pelotas (UFPEL) – agostini@inf.ufpel.edu.br

### 1. INTRODUÇÃO

Nos dias atuais, o consumo de vídeos é algo inerente ao cotidiano, seja para lazer, para aulas no formato remoto, para chamadas de vídeo e para várias outras aplicações. O consumo de vídeos é tão presente ao nosso cotidiano que de acordo com a Cisco (2023), cerca de 82% do tráfego total na internet é composto por consumo de vídeo.

Além disso, o Brasil atualmente é o quarto país da América Latina que mais consome esse tipo de mídia. De acordo com um estudo realizado pela Kantar Ibope Media (2022) sobre o tempo de consumo individual, os brasileiros consomem cerca de cinco horas e trinta e sete minutos por dia em vídeos digitais na internet, ficando apenas atrás da Argentina, Panamá e Chile, respectivamente.

Todavia, apesar da praticidade, a transmissão de vídeo consome muita banda da rede de comunicação. Como exemplo, um vídeo em *full high definition* (FHD), que possui uma resolução de 1920x1080 pixels, sem compressão, com 30 quadros por segundo e com 24 bits por pixel, consome cerca de 1,5 bilhões de bits por segundo (1,5 Gb), sendo que o razoável seria uma taxa de menos que 20 milhões de bits por segundo (20 Mb).

Além disso, de acordo com CISCO (2023), é esperado um aumento de 33% no consumo de vídeos *ultra high definition* (UHD) ou 4K, que tem existe uma taxa de transmissão quatro vezes maior do que vídeos em FHD (CISCO, 2023). Com todo esse cenário desenhado, é essencial diminuir a quantidade de bits necessárias para representar os vídeos digitais, mantendo a sua qualidade. Essa área de investigação é chamada de codificação de vídeos e a relação entre o número de bits e a qualidade do vídeo é chamada de eficiência de codificação. Assim, o cenário atual exige uma eficiência de codificação tão elevada quanto possível para viabilizar as aplicações futuras nessa área.

O grande desafio é que, para ampliar a eficiência de codificação, os codificadores atuais exigem um esforço computacional extremamente elevado.

Este trabalho busca desenvolver soluções para a redução de complexidade da compressão de vídeos 360°, a partir do codificador de vídeos *Versatile Video Coding* (VVC), que é o atual padrão estado-da-arte para a codificação de vídeos da ISO e da ITU-T. Vídeos 360° são um tipo de mídia cada vez mais popular e consiste na captura do vídeo em uma esfera que atinge 360° de longitude e 180° de latitude, onde o espectador se localiza no centro e pode optar para onde olhar. De acordo com Ye (2020, *apud* Rosado, 2022), para fornecer aos usuários uma qualidade de experiência satisfatória em uma sessão de visualização imersiva, o vídeo 360° requer uma largura de banda muito maior. Assim, eles requerem uma codificação eficiente para viabilizar seu uso de uma forma mais ampliada.

Quando o vídeo 360° é codificado, ele é transformado em um vídeo em duas dimensões a partir de uma projeção e a projeção mais usada é a chamada projeção equiretangular (ERP). Nesse processo de projeção, as regiões superiores e inferiores (regiões polares) do vídeo sofrem uma distorção de esticamento horizontal.

Esse trabalho visa reduzir o custo computacional da codificação de vídeos 360° com o VVC, visando o mínimo de perda da qualidade do vídeo. A ideia explorada neste trabalho é modificar a ferramenta chamada *Intra Subpartition* (ISP), aproveitando as características do esticamento horizontal que acontece nos polos da imagem por conta da projeção ERP. A ferramenta ISP divide o componente de luminância de um bloco vertical ou horizontal em subpartições com o mesmo tamanho. A hipótese investigada é que, por conta do esticamento horizontal nas regiões polares, provavelmente é possível deixar de usar a ferramenta ISP vertical nestas regiões, sem custo adicional em termos de eficiência de codificação.

## 2. METODOLOGIA

Para esse trabalho, visando explorar a deformação nas zonas polares, busca-se verificar se a desativação dos modos ISP verticais nestas zonas do vídeo causa reduções no custo computacional na codificação de vídeos 360°. Dessa forma, será explorada a observação de Rosado (2022) que, a partir da

análise de mapas de calor de uso da ferramenta ISP, verificou que os modos ISP horizontais são mais usados nas regiões polares, especialmente em tamanhos de blocos maiores.

Por tratar-se de uma pesquisa em caráter inicial, foram testados cinco quadros das sequências de vídeo definidas nas *Common Test Conditions* (CTC), que são as condições de teste padronizadas para o padrão VVC. Foi avaliada a hipótese de que ao desabilitar os modos verticais nas zonas polares poderá ser observado um melhor desempenho da codificação VVC para esse tipo de vídeo, com perdas de qualidade mínimas e ganhos em termos de esforço computacional.

O teste usou quatro vídeos: *Chairlift Ride*, *Harbor*, *Kite Flite* e *Trolley*. Cada um deste vídeos possui características distintas, seguindo assim, os padrões de teste da codificação de vídeos da *Join Video Exploration Team* (JVET).

Para avaliar a qualidade da codificação foi usada a métrica *Weighted-to-Spherically Peak Signal-to-Noise Ratio* (WS-PSNR), que é tipicamente utilizada quando da avaliação de soluções para vídeos 360°. Quanto maior o WS-PSNR, menos ruído foi inserido no vídeo e, conseqüentemente, a qualidade da imagem fica melhor. Esse método de avaliação mede a quantidade de ruído inserido no vídeo pelo processo de codificação, balanceando os valores de acordo com a deformação da zona do vídeo. Esta métrica é usada pois:

Dessa forma, quando a projeção ERP é utilizada, as amostras na região do equador que apresentam esticamento suave ou até mesmo nulo, tem um impacto mais significativo no cálculo da qualidade global, enquanto que as amostras nas regiões superior e inferior, onde o esticamento é mais severo, tem um impacto menos significativo (STORCH, 2020, p. 65).

### 3. RESULTADOS E DISCUSSÃO

Após a codificação, foi calculado o WS-PSNR dos vídeos e os dados foram os seguintes: *Harbor* teve um WS-PSNR de 46,9874 db antes da desativação dos modos ISP-Vertical e 47,0206 db após a modificação, um aumento de 0,07 db. O vídeo *Chairlift Ride* teve um WS-PSNR de 45,5120 db antes da modificação e de 45,5104 db após, uma queda de 0,003 db. O vídeo *Kite Flite* teve um WS-PSNR de 45,0347 db antes da modificação e 45,0359 db

após, com um aumento de 0,002 db. Por fim, o vídeo *Trolley* teve um WS-PSNR de 44,9972 db antes da modificação e 44,9945 db após, com uma queda de 0,006 db.

#### 4. CONCLUSÕES

Os resultados dos testes mostraram variações mínimas em termos de qualidade, com a desativação da ferramenta ISP vertical, validando a hipótese inicial do trabalho de que o modo vertical da ferramenta ISP é pouco relevante para as regiões planares dos vídeos. Os testes em *Harbor* e *Kite Flite* mostram um aumento no WS-PSNR, já *Chairlift Ride* e *Trolley* mostram uma queda no WS-PSNR. A variação de valores se dá, pois os vídeos do conjunto de CTCs possuem características distintas. Como foram testados poucos quadros dos vídeos, uma análise com a codificação completa de cada um deve se mostrar mais efetiva do que esses testes. A próxima etapa do trabalho é avaliar todos os quadros dos vídeos e também colher dados de tempo de execução para medir qual é o ganho em termos de esforço computacional, com a desativação do modo ISP vertical nas regiões polares dos vídeos 360°.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

CISCO. Cisco Annual Internet Report (2018–2023) White Paper. 2020. Disponível em: <<https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html>>. Acesso em: 16 de set. 2023.

Kantar Ibope Media. Inside video novos horizontes e descobertas. 2022. Disponível em: <<https://www.kantaribopemedia.com/wp-content/uploads/2022/05/Inside-Video-2022-Kantar-IBOPE-Media.pdf>>. Acesso em: 16 de set. 2023.

ROSADO, Rogério. Redução de Complexidade no Processo de Codificação de Vídeos 360° do Padrão VVC. 2022.

STORCH, Iago. Exploração das Distorções da Projeção ERP para Redução de Complexidade da Codificação de Vídeos Omnidirecionais. 2020. Disponível em: <[https://guaiaca.ufpel.edu.br/bitstream/handle/prefix/6086/Dissertacao\\_Iago\\_Storch.pdf?sequence=1&isAllowed=y](https://guaiaca.ufpel.edu.br/bitstream/handle/prefix/6086/Dissertacao_Iago_Storch.pdf?sequence=1&isAllowed=y)>. Acesso em: 17 de set. 2023.