

## ANÁLISE DE VIABILIDADE DE EXPLORAÇÃO DAS DISTORÇÕES ERP PARA REDUÇÃO DE COMPLEXIDADE NA CODIFICAÇÃO DE VÍDEOS OMNIDIRECIONAIS

**VITÓRIA MARCA SANTA LUCIA<sup>1</sup>; IAGO STORCH<sup>1</sup>; DANIEL MUNARI VILCHEZ PALOMINO<sup>1</sup>**

<sup>1</sup>Universidade Federal de Pelotas – (vmslucia, icstorch, dpalomino)@inf.ufpel.edu.br

### 1. INTRODUÇÃO

A crescente adoção de vídeos omnidirecionais impõe desafios específicos à codificação devido às distorções introduzidas pela projeção equiretangular (ERP), que permite representar a esfera em um plano retangular. Entretanto, essa projeção apresenta distorções geométricas, como por exemplo, o estiramento nas regiões polares, que afetam a eficiência de decisões de particionamento e predição dos codificadores modernos STORCH (2020). Ao analisar as decisões tomadas pelo software de codificação de vídeo, é constatado que, nas regiões inferior e superior do *frame*, observa-se uma maior tendência na ocorrência de blocos grandes e horizontais. Logo, há a oportunidade de simplificar e implementar heurísticas a fim de reduzir o tempo e a complexidade computacional na predição intra-quadros DOS SANTOS (2024).

Neste estudo, é utilizado o VVC (Versatile Video Coding), que foi finalizado em julho de 2020 e é o mais recente padrão internacional de codificação de vídeos. A Figura 1 ilustra a CTU (*Coding Tree Unit*), unidade básica de processamento do VVC de 128x128 pixels, uma estrutura hierárquica, encarregada de organizar particionamentos e determina os testes avaliados pelo codificador. Por esse motivo, intervenções direcionadas às decisões de particionamento em CTUs das regiões polares podem gerar ganhos de desempenho significativos. O principal desafio é garantir um *trade-off* favorável entre redução de tempo/complexidade e preservação da eficiência de compressão, avaliada aqui por métricas como BD-Rate e BD-WS-PSNR.

O presente trabalho tem como objetivo investigar a viabilidade de heurísticas aplicadas às zonas polares dos *frames* representados em ERP, focando em estratégias que limitem testes de particionamento e modos intra em CTUs dessas regiões. Posteriormente, pretende-se avaliar quantitativamente os *trade-offs*.

### 2. METODOLOGIA

No VVC, a CTU que pode ter tamanho máximo de 128×128 amostras de luminância. Cada CTU é recursivamente particionada em CUs (Coding Units) (Figura 1) segundo a estrutura QTMT (Quadtree + Multi-Type Tree), permitindo subdivisões quádruplas, binárias e ternárias (ex.: BTB/BTH/TTV/TTH) e particionamentos ISP, ilustrado na Figura 2. As CUs determinam os blocos sobre os quais são realizados os testes de predição intra-quadro. Intervir nas decisões de particionamento a nível de CTU/CU permite reduzir o conjunto de testes executados pelo codificador e, portanto, sua complexidade computacional.

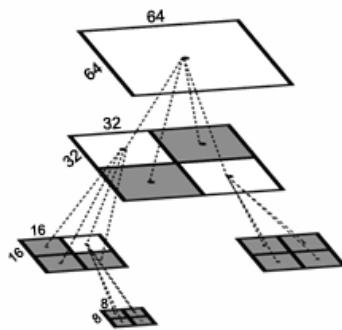


Figura 1 - Subdivisão de uma CTU em várias CUs. Adaptado de (CORREA et al., 2015)

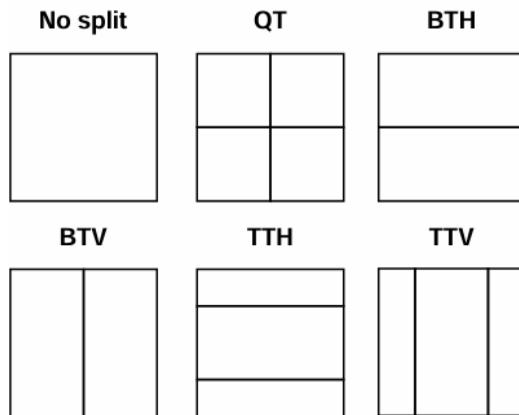


Figura 2 - Tipos de particionamento disponíveis na QTMT do VVC. Adaptado de (DOS SANTOS, 2024)

No presente trabalho, foi utilizado o *software* VTM, uma implementação de referência do padrão VVC para realização das análises. No *software*, foi integrada a biblioteca 360Lib, uma biblioteca criada pelos mesmos desenvolvedores do VVC para adicionar o suporte a vídeos 360°, bem como uma rotina de *logging* que registra em um arquivo .csv, para cada CTU, os particionamentos que serão **testados** pelo codificador. Em paralelo, os particionamentos julgados ótimos pelo *software* serão realmente **usados** para a codificação e ficarão armazenados no *bitstream* do *frame* codificado. Essa diferenciação permite comparar os comportamentos e identificar quais particionamentos, em quais regiões do *frame*, estão sendo demasiadamente testadas e pouco utilizadas.

Foram utilizados 4 vídeos omnidirecionais e 4 vídeos convencionais, com o objetivo de também poder comparar o comportamento específico dos vídeos 360°. Os vídeos utilizados, definidos pelas CTCs360 (Condições Comuns de Teste para Vídeos Omnidirecionais) foram: *ChairliftRide*, *Harbor*, *Skateboard In Lot* e *Trolley*, todos com resolução 8192x4096. Já para as CTCs (para vídeos convencionais), foram escolhidos os vídeos *Campfire*, *CatRobot*, *ParkRunning3* e *Tango2*, todos com resolução 3840x2160 (BOYCE; ALSHINA; YE, 2017). Além disso, foram empregados quatro QPs: 22, 27, 32 e 37, conforme definido pelas CTCs360 e CTCs.

Foram codificados 10 *frames* de cada vídeo, para cada QP. Após a realização dos experimentos, os dados foram processados através de *scripts*. Os dados foram agrupados a cada 128 pixels, denominados Linha de CTU, para ser possível analisar em qual região do *frame* as ações estão sendo tomadas pelo codificador. Também foi calculado o *aspect ratio*, convertendo os tamanhos de bloco possíveis dentro das CUs em números, a partir do cálculo de largura dividido pela altura. Um bloco de tamanho 64x16 teria o *aspect ratio* equivalente a 4, por exemplo.. No processamento de dados, foi calculada a média por QP de todos estes parâmetros, de tal forma que fossem gerados gráficos para cada QP, com a distribuição de cada *aspect ratio* ao longo das Linhas de CTU.

### 3. RESULTADOS E DISCUSSÃO

Os experimentos geraram gráficos que relacionam, para cada linha de CTU, a distribuição de tamanhos de blocos testados e efetivamente usados, tanto em vídeos omnidirecionais (ERP) quanto em sequências convencionais, nos quatro QPs avaliados (22, 27, 32 e 37). A normalização aplicada aos dados (cada valor dividido pelo somatório do *dataset* e ajuste de escala pelo valor máximo) permitiu comparar os resultados em diferentes sequências e resoluções sob a mesma base de escala.

Nos **vídeos omnidirecionais**, observa-se que as bandas polares apresentam maior predominância de blocos horizontais, enquanto a banda central exibe maior diversidade de particionamentos. Esse padrão confirma a hipótese de que a distorção da projeção ERP (esticamento horizontal) influencia a escolha dos modos de particionamento.

Nos **vídeos convencionais**, em contrapartida, as distribuições mostraram-se mais homogêneas entre linhas de CTU, sem concentração marcante de blocos em regiões específicas do quadro. Essa diferença reforça o caráter singular das projeções ERP, validando a motivação do estudo.

Os gráficos de “teste” (blocos avaliados pelo codificador) e “codificado” (blocos efetivamente escolhidos) evidenciam redundância nas regiões polares das sequências 360°. Diversos tamanhos e particionamentos são testados nessas regiões, mas poucos são efetivamente utilizados. Esse achado confirma a possibilidade de aplicar heurísticas regionais que limitem os testes nessas zonas, com impacto mínimo na eficiência de compressão.

A Figura 3 ilustra esse comportamento para o QP 22 em vídeos omnidirecionais. Nota-se que blocos verticais (*aspect ratio* < 1) são frequentemente testados, mas raramente selecionados pelo codificador. Em contrapartida, blocos horizontais apresentam uso proporcionalmente maior em relação ao número de vezes em que são testados. Esse resultado reforça a pertinência de heurísticas que reduzam ou eliminem testes de partições verticais nas regiões polares.

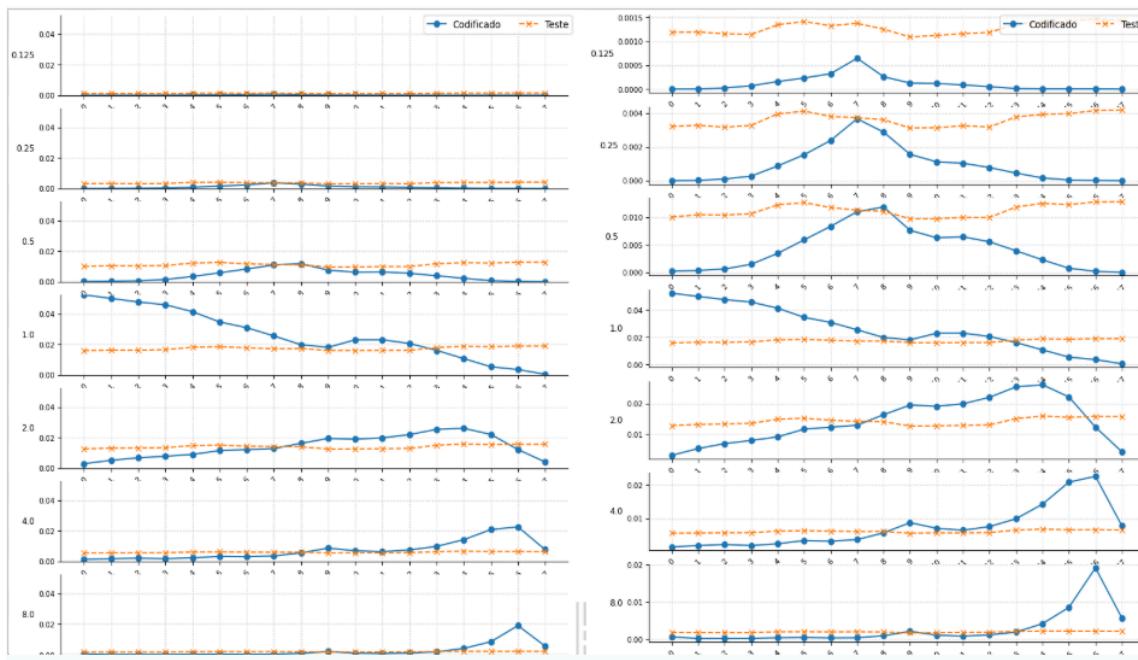


Figura 3 - Gráfico dos resultados do QP 22, dados ponderados, de vídeos omnidirecionais.

#### 4. CONCLUSÕES

O presente trabalho validou as afirmações acerca do comportamento ímpar dos vídeos omnidirecionais feitas por STORCH (2020) e DOS SANTOS (2024). Conclui-se que há espaço para a implementação de heurísticas para reduzir o número de testes feitos na etapa de codificação da predição intra-quadros do VVC, visto que o comportamento desse tipo de vídeo, por conta da projeção ERP, é singular.

A perspectiva é implementar heurísticas que não apenas aliminem o número de testes de blocos verticais nas zonas polares dos frames, mas sim reduzam a quantidade de vezes que esse perfil de bloco é escolhido comprometendo o mínimo possível métricas como o BD-Rate e BD-WS-PSNR.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

- BOYCE, J.; ALSHINA, E.; ABBAS, A.; YE, Y. JVET common test conditions and evaluation procedures for 360 video. **JVET-H1030**, Chengdu, CN, [S.I.], 2017.
- CORREA, G.; ASSUNCAO, P. A.; AGOSTINI, L. V.; SILVA CRUZ, L. A. da. Fast HEVC Encoding Decisions Using Data Mining. **IEEE Transactions on Circuits and Systems for Video Technology**, [S.I.], v.25, n.4, p.660–673, April 2015.
- DOS SANTOS, O. F. **Heurísticas para Reduzir o Custo Computacional da Codificação de Vídeos Omnidirecionais na Predição Intra-Quadro do Padrão VVC**. 2024. 40f. Trabalho de Conclusão de Curso do Bacharelado em Ciência da Computação — Universidade Federal de Pelotas.
- STORCH, I. C. **Exploração das Distorções da Projeção ERP para Redução de Complexidade da Codificação de Vídeos Omnidirecionais**. 2020. 175p. Mestrado em Ciência da Computação — Universidade Federal de Pelotas, Pelotas.