

Heurística de Redução de Tempo de Codificação para a Predição Intra-Quadro do Padrão VVC com Foco em Vídeos 360°

FRANKLIN SALES DE OLIVEIRA¹; OTÁVIO FERNANDES DOS SANTOS²;
LUCAS SUPERTI DA SILVA³; VANESSA ALDRIGHI⁴; ROGÉRIO CASSANTA
ROSADO⁵; LUCIANO VOLCAN AGOSTINI⁶

¹Universidade Federal de Pelotas (UFPel) – fsoliveira@inf.ufpel.edu.br

²Universidade Federal de Pelotas (UFPel) – otavio.fsanto@inf.ufpel.edu.br

³Universidade Federal de Pelotas (UFPel) – lsdsilva@inf.ufpel.edu.br

⁴Universidade Federal de Pelotas (UFPel) – vanessa.a@inf.ufpel.edu.br

⁵Universidade Federal de Pelotas (UFPel) – rcassanta@inf.ufpel.edu.br

⁶Universidade Federal de Pelotas (UFPel) – agostini@inf.ufpel.edu.br

1. INTRODUÇÃO

A demanda por vídeos imersivos, que proporcionam ao espectador uma experiência envolvente e interativa, tem crescido significativamente nos últimos anos, com aplicações que vão do entretenimento à educação (STORCH, 2020). Os vídeos de 360° ou omnidirecionais, criam a sensação de imersão ao permitir que o usuário explore o ambiente em todas as direções, seja movendo a câmera em dispositivos móveis ou utilizando tecnologias como óculos de realidade virtual. No entanto, essa riqueza de conteúdo visual requer uma enorme quantidade de dados, intensificando a necessidade de métodos eficientes de codificação e compressão (SANTOS, 2024).

Os vídeos de 360° são capturados na forma de uma esfera, com múltiplas câmeras convencionais, cujos conteúdos são agrupados para gerar a sensação de imagem contínua e imersiva. Para a codificação dos vídeos 360°, a imagem esférica é projetada em um plano 2D e um codificador de vídeo convencional é aplicado. Existem diversas formas de projeção, mas a mais usada, e foco deste trabalho, é a projeção Equirectangular (ERP), que é a mesma projeção mais usada nos mapas que representam o globo terrestre.

De modo geral, um padrão de codificação convencional aproveita características dos vídeos, como redundâncias de dados espaciais e temporais, para alcançar altas taxas de compressão, mas, para tanto, o custo computacional é bastante elevado. Para atender as demandas emergentes na área de codificação de vídeos, surgiu o padrão *Versatile Video Coding* (VVC), que traz uma solução avançada em codificação de vídeos, oferecendo compressão mais eficiente. Contudo, o VVC tem um custo computacional muito mais elevado do que os padrões anteriores, especialmente ao lidar com vídeos 360°, que requerem altas resoluções e geram uma maior carga de trabalho na codificação (STORCH, 2020).

O padrão VVC segmenta cada frame do vídeo em blocos menores, chamados de Coding Tree Units (CTUs) que são processados de forma mais eficiente, cobrindo uma área quadrada de até 128x128 pixels. CTUs podem ser recursivamente divididos em blocos ainda menores chamados de Coding Units (CUs). Esse processo de segmentação segue uma estrutura conhecida como Quadtree with nested multi-type tree (QTMT). Uma CU pode ser considerada indivisível, realizando-se a codificação com seu tamanho atual. Alternativamente, uma CU pode ser segmentada por meio de partições quaternárias (QT), binárias

(BT) ou ternárias (TT). A estrutura QT efetua a divisão de uma CU em quatro sub-CUs quadradas de tamanhos iguais. Por sua vez, a estrutura BT pode dividir a CU simetricamente, seja de forma horizontal (BTH) ou vertical (BTV), resultando em duas sub-CUs. Já a estrutura TT realiza uma divisão da CU em três sub-CUs nas proporções de 1:2:1, podendo essa divisão ocorrer nas direções horizontal (TTH) ou vertical (TTV) (SANTOS, 2024).

Além da estrutura de partição inovadora, o VVC introduziu várias novas ferramentas ou melhoramentos expressivos nas ferramentas anteriores, nas diversas etapas da codificação. Na predição intra-quadro, que é foco deste trabalho, várias ferramentas foram ampliadas ou incluídas. As ferramentas de interesse deste trabalho são a *Angle Intra Prediction* (AIP) e *Intra Subpartition* (ISP). Na ferramenta AIP, a CU atual é predita usando interpolações das amostras acima e à esquerda das CUs vizinhas em diversos ângulos. No VVC os modos angulares foram ampliados para 65 (eram 33 em padrões anteriores). A ferramenta ISP é uma novidade do VVC e subdivide as CUs em partes menores na horizontal ou vertical, permitindo que a predição e a codificação sejam realizadas separadamente em cada subpartição (PFAFF *et al*, 2021).

Este trabalho apresenta uma pesquisa realizada pelos integrantes do *Video Technology Research Group (ViTech)*, da Universidade Federal de Pelotas, que visa explorar os efeitos das distorções na imagem em projeção ERP de vídeos 360° sobre a predição intra-quadro do VVC. O objetivo é reduzir o tempo de execução com penalidades mínimas na eficiência de codificação. Mais especificamente, o trabalho apresenta uma heurística que explora os esticamentos horizontais presentes nas regiões superior e inferior das imagens com projeção ERP. Assim, nestas regiões, as CUs verticais (BTV e TTV) são desabilitadas, os modos angulares mais verticais da AIP não são testados e os subparticionamentos verticais da ISP não são avaliados.

2. METODOLOGIA

A princípio, cada quadro do vídeo foi segmentado em três bandas, de acordo com a Figura 1: banda superior, banda central e banda inferior, conforme a metodologia proposta por (STORCH, 2020). A banda central contém 50% das amostras do quadro, enquanto as bandas superior e inferior possuem, cada uma, 25% das amostras. É importante notar que na projeção ERP o equador é mapeado com poucas distorções, enquanto os pólos são esticados horizontalmente (STORCH, 2020).

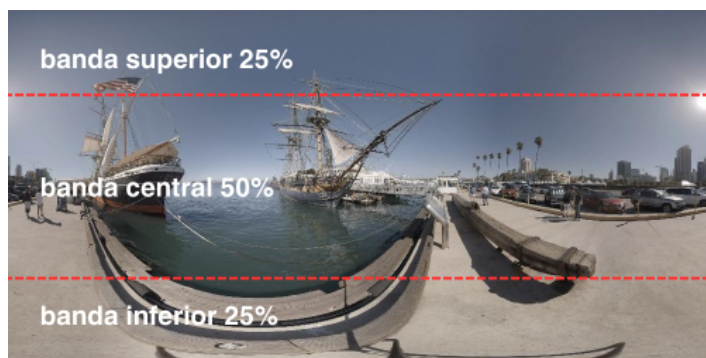


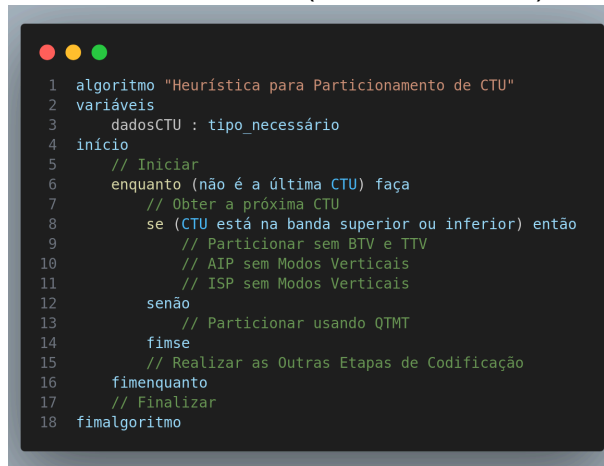
Figura 1 – Divisão do quadro em bandas. Fonte: (SANTOS, 2024).

Em todos os experimentos, foi utilizado o software de referência do padrão VVC, conhecido como *VVC Test Model* (VTM), na versão 19.0, juntamente com a biblioteca 360Lib, versão 13.4, essencial para trabalhar com projeções em formato ERPI.

A base experimental seguiu as Condições Comuns de Teste (CTC) do VVC para vídeos omnidirecionais. A configuração escolhida foi a All-Intra, o que significa que todos os quadros são codificados exclusivamente pela predição intra-quadro, que é o foco deste trabalho (SANTOS, 2024). Os parâmetros de quantização (QP) foram 22, 27, 32 e 37 e um total de oito sequências de vídeo foram utilizadas, seguindo as definições das CTCs para vídeos 360°. Assim, foram realizadas 64 codificações, metade para a versão original do VTM e metade para a versão do VTM com a heurística incluída. Cada sequência em cada versão do codificador foi codificada quatro vezes, uma para cada QP.

Foram utilizadas duas métricas para avaliar os impactos na eficiência de codificação: *Bjontegaard Delta Bitrate* (BD-RATE) e *Bjontegaard Delta Weighted-to-Spherically Uniform PSNR* (BD-WSPSNR). O BD-RATE possibilita a comparação entre dois codificadores, mostrando como a taxa de bits varia para manter a mesma qualidade do vídeo. O BD-WSPSNR, por sua vez, analisa como a qualidade objetiva é afetada a uma taxa de bits fixa (SANTOS, 2024). A métrica usada para avaliar os ganhos em custo computacional foi a redução percentual do tempo de execução com e sem o uso da heurística.

A Figura 2 apresenta o pseudo algoritmo da heurística proposta. Inicialmente, é preciso analisar a posição de cada CTU para verificar se a CTU está localizada na banda superior ou inferior do quadro. Caso não esteja, a codificação continua normalmente. No entanto, se estiver, a heurística é aplicada. Nesse caso as partições BTV e TTV são desabilitadas e as ferramentas AIP e ISP são executadas sem os modos verticais (SANTOS, 2024).



```
1 algoritmo "Heurística para Particionamento de CTU"
2 variáveis
3   dadosCTU : tipo_necessário
4 início
5   // Iniciar
6   enquanto (não é a última CTU) faça
7     // Obter a próxima CTU
8     se (CTU está na banda superior ou inferior) então
9       // Particionar sem BTV e TTV
10      // AIP sem Modos Verticais
11      // ISP sem Modos Verticais
12     senão
13       // Particionar usando QTMT
14     fimse
15     // Realizar as Outras Etapas de Codificação
16   fimenquanto
17   // Finalizar
18 fimalgoritmo
```

Figura 2 – Pseudo algoritmo da heurística proposta.

3. RESULTADOS E DISCUSSÃO

Os resultados obtidos estão apresentados na Tabela 1, onde são apresentados os resultados médios para cada sequência (para os quatro QPs) e também os resultados médios gerais. Estes resultados indicam uma redução média de 8,27% no tempo total de codificação, com uma perda insignificante em eficiência de codificação. A destacar que o ganho de tempo considera todo o VVC, incluindo o restante da predição intra-quadro e as etapas de transformadas, quantização, codificação de entropia e filtros de laço.

| Sequência | BD-Rate (%) | BD-WSPNSR (dB) | Economia de tempo (%) |
|-----------------|-------------|----------------|-----------------------|
| ChairliftRide | 0,46 | -0,017 | 9,6 |
| Gaslamp | 0,70 | -0,032 | 7,6 |
| Harbor | 0,63 | -0,030 | 10,4 |
| KiteFlite | 0,20 | -0,012 | 10,1 |
| SkateboardInLot | 0,18 | -0,007 | 9,4 |
| SkateboardTrick | 0,25 | -0,012 | 2,2 |
| Train | 0,16 | -0,008 | 6,2 |
| Trolley | 0,11 | -0,006 | 10,7 |
| Média | 0,34 | -0,016 | 8,27 |

Tabela 1 – Resultados com a Heurística. Fonte: (SANTOS, 2024).

4. CONCLUSÕES

Este trabalho apresentou uma heurística para reduzir o custo computacional da predição intra-quadro em vídeos 360° codificados no padrão VVC. A heurística aproveita as características da projeção equiretangular para descartar blocos verticais e modos de predição verticais nas bandas superior e inferior dos quadros, acelerando o processo de codificação, sem comprometer a sua eficiência. Comparado a outros métodos, nossa abordagem atinge um melhor equilíbrio entre a redução de tempo e a preservação da eficiência de codificação.

5. REFERÊNCIAS BIBLIOGRÁFICAS

CORREA, G.; ASSUNÇÃO, 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*, Nova York, v.25, n.4, p.660-673, 2015.

PFAFF, J. *et al.* Intra Prediction and Mode Coding in VVC. *IEEE Transactions on Circuits and Systems for Video Technology*, Nova York, v.31, n.10, p.3834-3847, 2021.

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 (Bacharel em Ciência da Computação) - Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas.

STORCH, I.C. **Exploração das distorções da projeção ERP para redução de complexidade da codificação de vídeos omnidirecionais**. 2020. 175f. Dissertação (Mestrado em Ciência da Computação) - Programa de Pós-Graduação em Computação, Universidade Federal de Pelotas.