

Avaliação de Resiliência a Erros de Armazenamento Aproximado na Intra Predição de Decodificadores VVC

MATHEUS ISQUIERDO¹; RENIRA SOARES²; FELIPE SAMPAIO³; BRUNO
ZATT⁴; DANIEL PALOMINO⁵

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

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

³Instituto Federal do Rio Grande do Sul – felipe.sampaio@farroupilha.ifrs.edu.br

⁴Universidade Federal de Pelotas – zatt@inf.ufpel.edu.br

⁵Universidade Federal de Pelotas – dpalomino@inf.ufpel.edu.br

1. INTRODUÇÃO

Ao longo dos anos, tem crescido expressivamente tanto a produção quanto o consumo de vídeos de alta resolução (HD), conforme aumenta a demanda por taxas de quadros e resoluções maiores, como 60 e 120 quadros por segundo e vídeos Ultra-HD (4K e 8K). Esse crescimento se deve em parte pela popularização de dispositivos móveis, responsáveis por 70% do tempo de exibição do YouTube (GOOGLE, 2018) e consideravelmente mais limitados em termos de capacidade computacional e energética. Isso evidencia a necessidade de pesquisa sobre sistemas de codificação de vídeo com o foco em aperfeiçoar o armazenamento e transmissão de tal quantidade de dados.

Com vista nisso, em 2020, o *Joint Video Experts Team* (JVET) finalizou o desenvolvimento do *Versatile Video Coding* (VVC) (BROSS et al., 2021), alcançando até 40% de redução na taxa de bits necessária para representar vídeos com a mesma qualidade que o *High Efficiency Video Coding* (HEVC), seu predecessor (WIECKOWSKI et al., 2021). Entretanto, tais ganhos só foram possíveis com o aumento da complexidade computacional. Por exemplo, decodificadores VVC apresentam 1,5 e 1,8 vezes a complexidade computacional do HEVC para *Random Access* e *All Intra*, respectivamente, enquanto requisitos de largura de banda de memória aumentaram 3 vezes (PAKDAMAN et al., 2020).

Muitos dos avanços no VVC foram na intra predição, como o dobro de modos angulares em comparação ao HEVC, entre outros. Com comunicação de memória sendo um limitante conhecido na computação (CERVEIRA et al., 2020), estratégias de redução em sistemas de codificação de vídeo são geralmente adotadas mirando em reduzir o consumo energético (TEIMOORI et al., 2018). Trabalhos que implementaram sistemas dedicados para o HEVC, entre outros, geralmente consideram uma memória dedicada para armazenar pixels vizinhos usados na intra predição (PALOMINO et al., 2012), denominada *Buffer* de Linha de Referência (BLR). Nessa linha, armazenamento aproximado é um paradigma promissor para melhorar tempo de acesso, consumo energético, entre outros. Tais ganhos são obtidos ao custo de precisão dos dados armazenados e, conseqüentemente, de qualidade dos resultados, que, entretanto, podem ser aceitáveis em aplicações resilientes, como mineração de dados, processamento de imagens, aprendizado de máquina, entre outros (TEIMOORI et al., 2018).

Esse trabalho apresenta uma avaliação de resiliência a erros do passo de intra predição dos decodificadores VVC quando armazenamento aproximado é aplicado ao *buffer* BLR. Primeiro, desenvolvemos uma ferramenta para simular armazenamento aproximado com valores típicos da literatura e, então, avaliamos os efeitos na qualidade da imagem final dos vídeos decodificados com diferentes configurações no *buffer* BLR para diferentes cenários.

2. METODOLOGIA

Para avaliar o impacto da memória aproximada no *buffer* BLR durante a intra predição no decodificador do VVC, criamos um *framework* com duas etapas principais. Primeiro, usando a implementação VVC *Test Model* (VTM) na sua versão 14.0 (JVET, 2021), 8 vídeos foram selecionados arbitrariamente das Condições Comuns de Teste (CCT) e codificados com base nelas, usando Parâmetros de Quantização (PQ) de 22, 27, 32 e 37 para cada vídeo, gerando 32 arquivos binários. A configuração usada para este trabalho foi a *Random Access* (RA), onde o vídeo é codificado em conjuntos de 33 quadros, um período intra composto por dois Grupos de Imagens (DGIs). Nele, os 2 quadros temporalmente mais externos (quadros-I) são codificados apenas com intra predição e servem de referência para os demais (quadros-B), que podem também fazer uso de inter predição e seguem uma hierarquia com dependências em níveis entre eles. Apenas o primeiro período intra de cada vídeo foi codificado.

Em seguida, para a decodificação, usamos um VTM 14.0 modificado para rodar junto à Pin 3.19 (INTEL, 2021), uma plataforma de instrumentalização binária, e uma ferramenta customizada, chamada *Pintool*, para injeção de erros na forma de inversões aleatórias de bits. Essa *Pintool* inspeciona instruções em nível de processador, verifica se elas lêem a memória nos intervalos determinados do BLR e, caso sim, faz a injeção de erros, com uma frequência definida por Taxas de Erros de Bits (TEBs). Neste trabalho, a TEB usada foi a de $10E-03$. Erros não foram inseridos nos I-frames para obter uma visão mais limpa sobre a propagação de erros na hierarquia.

Devido a aleatoriedade dos erros, decodificamos cada arquivo binário 100 vezes para obtermos valores mais confiáveis, menos afetados por valores atípicos. A resiliência a erros é medida considerando a degradação da qualidade da imagem, medida em *Peak Signal-to-Noise Ratio* (PSNR) usando o vídeo original como referência. A perda de qualidade é calculada subtraindo o PSNR preciso, obtido pela decodificação do vídeo sem a injeção de erros, pelo PSNR aproximado, obtido da decodificação com inserção de erros no *buffer* BLR durante a intra predição. Para o cálculo do PSNR final de cada vídeo, os quadros-I não foram considerados, já que não sofreram injeção.

3. RESULTADOS E DISCUSSÃO

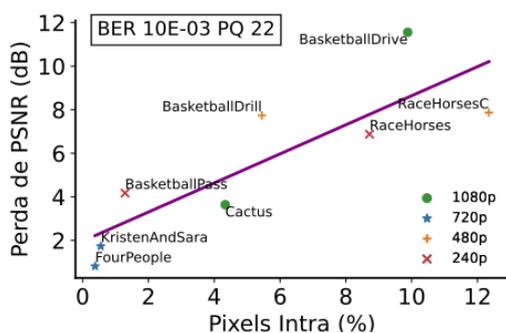


Figura 1: Correlação entre a prevalência de pixels intra em quadros-B e a perda de PSNR para a configuração RA

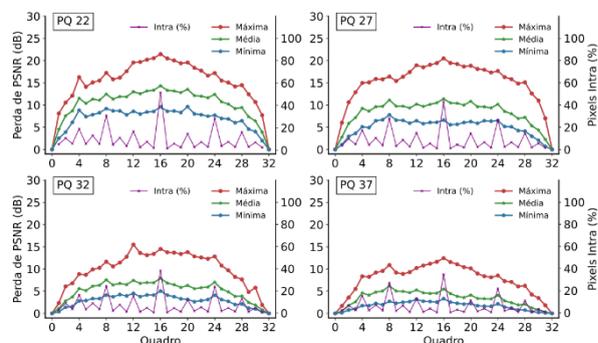


Figura 2: Perda de PSNR e prevalência de pixels intra no vídeo BasketballDrive com TEB de $10E-03$ para um período intra em RA

Análise 1: A Fig. 1 mostra um gráfico de dispersão com a correlação entre a prevalência de pixels decodificados pela intra predição com um *buffer* BLR aproximado e a perda de PSNR, considerando uma configuração RA e uma TEB de $10E-03$. Ela mostra a presença de uma correlação positiva entre os dois fatores, com um coeficiente de correlação Pearson de $+0,85$ para o PQ 22. Isso se deve pois conforme mais pixels são decodificados com o uso de armazenamento aproximado, mais erros são inseridos durante o processo. Por exemplo, é facilmente notável que a sequência *FourPeople*, a de menores perdas de PSNR na Fig. 1, também apresenta a menor prevalência de pixels decodificados com intra predição. Por outro lado, *BasketballDrive* possui a maior perda de PSNR tendo um dos maiores usos de intra predição. Esse comportamento se repete para todos os PQs, mesmo que em proporções diferentes.

Análise 2: A Fig. 2 apresenta uma análise quadro a quadro das perdas de PSNR e da prevalência de pixels intra do primeiro período intra do vídeo *BasketballDrive* codificado com a configuração RA e decodificado com um TEB de $10E-03$, para todos os PQs. Ela mostra as perdas de PSNR máxima, média e mínima (linhas vermelha, verde e azul, respectivamente) ao lado da porcentagem de pixels intra (linha roxa) para cada quadro-B decodificado, ignorando quadros-I. Nela, o principal a ser percebido é como as perdas de PSNR se correlacionam com a hierarquia do DGI. A primeira informação a ser notada é que o comportamento da perda de PSNR se assemelha a uma função quadrática negativa, com os maiores valores sendo atingidos no meio do período intra, entre GDIs. Tal comportamento pode ser compreendido por alguns motivos. Primeiro, ele é fortemente correlacionado com a prevalência de pixels intra nos quadros-B. Como visto na linha roxa, a presença de pixels intra é maior no quadro central do período intra. Ele também é regular para todos os PQs, devido ao referenciamento entre quadros-B na hierarquia e à ordem em que eles são codificados. Com mais pixels intra para serem decodificados com o *buffer* BLR aproximado, mais erros são inseridos no vídeo. O segundo motivo é a distância entre os quadros-B e os quadros-I, já que o codificador tende a achar as melhores referências temporais de um quadro em seus vizinhos mais próximos. No decodificador, com o *buffer* BLR aproximado, erros inseridos nos quadros-B em níveis mais baixos do GDI criam um efeito de propagação de erros, pois esses quadros-B afetados são usados como referência da decodificação dos quadros-B em níveis mais altos. Esse efeito é mais claro nos quatro primeiros (ou últimos) quadros. Conforme os quadros-B decodificados se afastam do quadro-I mais próximo, as perdas de PSNR rapidamente crescem, pois torna-se mais provável o uso de referências com erros.

4. CONCLUSÕES

Esse trabalho apresentou uma avaliação de resiliência a erros da etapa de intra predição dos decodificadores do padrão VVC sob a configuração *Random Access* com a aplicação de armazenamento aproximado no *buffer* BLR. Os resultados mostram maiores perdas na qualidade da imagem em vídeos e quadros que fazem um maior uso da intra predição em quadros-B, assim como um aumento de perdas conforme os quadros-B se afastam dos quadros-I, usados como referência.

5. REFERÊNCIAS BIBLIOGRÁFICAS

GOOGLE. **Global youtube mobile watch time statistics**. Think with Google, San Bruno, jun. 2018. Acessado em 18 jul. 2022. Online. Disponível em: <https://www.thinkwithgoogle.com/marketingstrategies/search/global-youtube-mobile-watch-time-statistics/>

BROSS, B.; CHEN, J. OHM.; J. R.; SULLIVAN, G. J.; WANG, Y. K. "Developments in International Video Coding Standardization After AVC, With an Overview of Versatile Video Coding (VVC)". **Proceedings of the IEEE**, vol. 109, no. 9, p. 1463-1493, 2021.

WIECKOWSKI, A.; LEHMANN, C.; BROSS, B.; MARPE, D.; BIATEK, T.; RAULET, M.; FEUVRE, J. L. A Complete End to End Open Source Toolchain for the Versatile Video Coding (VVC) Standard. **Proceedings of the 29th ACM International Conference on Multimedia**. New York, p. 3795–3798, 2021.

PAKDAMAN, F.; ADELIMANESH, M. A.; GABBOUJ, M.; HASHEMI, M. R. "Complexity analysis of next-generation vvc encoding and decoding". **IEEE International Conference on Image Processing (ICIP)**. IEEE, p. 3134–3138, 2020.

PALOMINO, D.; SAMPAIO, F.; AGOSTINI, L.; BAMPI, S.; SUSIN, A. "A memory aware and multiplierless vlsi architecture for the complete intra prediction of the hevc emerging standard". **19th IEEE International Conference on Image Processing (ICIP)**. IEEE, p. 201–204, 2012.

CERVEIRA, A.; AGOSTINI, L.; ZATT, B.; SAMPAIO, F. "Memory assessment of versatile video coding". **IEEE International Conference on Image Processing (ICIP)**. IEEE, p. 1186–1190, 2020.

TEIMOORI, M. T.; HANIF, M. A.; EJLALI, A.; SHAFIQUE, M. "Adam: Adaptive approximation management for the non-volatile memory hierarchies". **Design, Automation Test in Europe Conference Exhibition (DATE)**. p. 785–790, 2018.

JVET. **Vvcsoftware vtm. ver. 14.0**. Fraunhofer, Online, 26 ago. 2021. Acessado em 13 out. 2021. Online. Disponível em: <https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftwareVTM/-/tree/VTM14.0>

INTEL. **Pin - A Dynamic Binary Instrumentation Tool**. Intel Corporation, Santa Clara, 1 mai. 2021. Acessado em 13 out. 2021. Online. Disponível em: <http://software.intel.com/content/www/us/en/develop/articles/pin-a-dynamic-binary-instrumentation-tool.html>