

AVALIAÇÃO DA INTENSIDADE DE ACESSOS À MEMÓRIA NA COMPENSAÇÃO DE MOVIMENTO DO DECODIFICADOR VVC

MATHEUS ISQUIERDO¹; BRUNO ZATT²; DANIEL PALOMINO³

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

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

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

1. INTRODUÇÃO

Ao longo dos anos, o número de dispositivos móveis capazes de lidar com vídeos, seja em captura, manipulação, armazenamento e transmissão, tem crescido de forma expressiva, assim como a demanda por taxas de quadros e resoluções maiores, como 60 e 120 quadros por segundo (qps) e UHD (*Ultra High Definition*) 4k e 8K. Atualmente, mais de dois terços do tempo de exibição de vídeos do *Youtube* provém de dispositivos móveis (YOUTUBE, 2018). Com base nisso, dois desafios se fazem presentes: o desenvolvimento de tecnologias de codificação que diminuam a quantidade de dados necessária para representar esses vídeos, e técnicas para a redução da energia necessária para trabalhar com essas tecnologias de codificação.

Para tratar o primeiro desafio, a *Joint Video Experts Team* (JVET) lançou uma chamada de propostas, em 2017, com o objetivo de coletar ideias de ferramentas de codificação de vídeo de maior taxa de compressão em comparação com o *High Efficiency Video Coding* (HEVC), o padrão de codificação estado-da-arte até o momento (JVET, 2021). Em 2018 o projeto de padronização do *Versatile Video Coding* (VVC) foi iniciado e, em 2020, foi finalizado, gerando uma redução na taxa de bits de aproximadamente 30% (ZHANG et al., 2020), sem perda significativa de qualidade visual em relação ao HEVC. Entretanto, com o ônus do aumento de complexidade de 5 e 1,5 vezes para a codificação e decodificação, respectivamente (PAKDAMAN et al., 2020). Isso é consequência de técnicas mais complexas e computacionalmente intensivas em relação à memória, o que se caracteriza como um problema quando ela representa uma grande porção do consumo energético de sistemas computacionais.

Mesmo que o decodificador seja consideravelmente mais simples que o codificador, ele ainda se apresenta como um desafio, especialmente com o avanço da visualização de vídeos em dispositivos móveis, que contam com uma capacidade energética limitada. Um dos módulos responsáveis por essa complexidade é o de Compensação de Movimento, sendo responsável por, em média, 20% da complexidade total do decodificador, atrás apenas do Filtro de Deblocagem (PAKDAMAN et al., 2020).

A Compensação de Movimento faz a reconstrução de um quadro com base nos quadros já processados, chamados de Quadros de Referência (QDR). Os QDR são organizados em Listas de Referências (LDR), que em sistemas computacionais vão estar armazenados em memória. Sendo assim, é importante quantificar a quantidade de acessos à memória realizada pelo módulo de compensação de movimento para ter um entendimento detalhado que vise soluções de redução de consumo de energia.

Portanto, o principal objetivo deste trabalho é quantificar a intensidade dos acessos à memória dos QDR durante da Compensação de Movimento do decodificador do VVC. Como resultado, é esperado entender de forma mais clara

a quantidade de acessos realizada pelo módulo de estimativa de movimento para viabilizar futuras soluções de memória que aperfeiçoem a eficiência energética no decodificador VVC.

2. METODOLOGIA

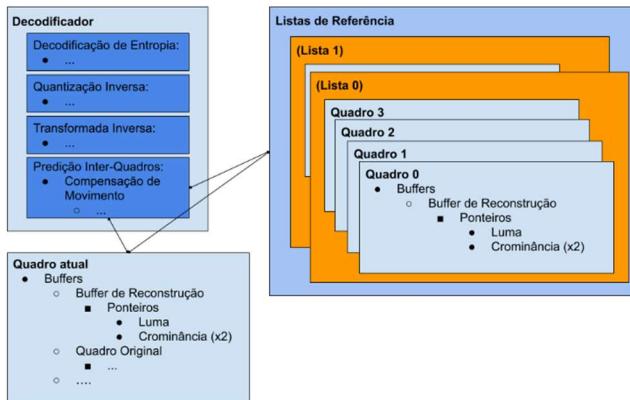


Figura 1: Esquema básico do Decodificador, ressaltando as Listas de Referência.

A Figura 1 apresenta um esquema superficial do decodificador VVC. Nele, o *bitstream* é processado pela Decodificação de Entropia. Os resíduos são gerados, pelas Quantização e Transformada Inversas, e somados a blocos presentes nos QDR das Listas de Referências durante a Compensação de Movimento. Neste experimento, quantificamos os acessos aos Buffers de Reconstrução dos quadros das Listas de Referência, para determinarmos a vazão de dados necessária para que se tenha um sistema em tempo real para diferentes resoluções de vídeo. Além disso, foi alterada a quantidade de QDR para a avaliação dos efeitos na quantidade de dados acessados.

O software VTM (*VVC Test Model*) 12.0, que contém uma implementação do codificador e do decodificador VVC, foi usado para a quantificação dos acessos. Todas as codificações e decodificações foram executadas de acordo com as condições comuns de teste determinadas pelo JVET e adotadas pela comunidade de codificação de vídeo (OHM, 2021). Três configurações de codificação foram escolhidas: Low Delay com quatro QDR (padrão) (LD 4), Low Delay com dois QDR (LD 2) e Low Delay com um QDR (LD 1).

A avaliação de acessos à memória durante a decodificação foi feita utilizando a Pin 3.19, uma plataforma de *profiling* e instrumentalização binária da Intel (INTEL, 2021). Ela foi associada a uma ferramenta customizada (*pintool*) para contabilizar acessos feitos aos QDR durante funções específicas, neste caso, os referentes à Compensação de Movimento, além de identificar durante a decodificação de qual quadro ativo esses acessos foram feitos.

As taxas de transferência são calculadas com base no número e largura dos acessos, número de quadros referenciados e a taxa de quadros por segundos (qps). As taxas de transferência foram normalizadas para 30 qps, pois os vídeos utilizados têm taxas de quadros diferentes entre si.

3. RESULTADOS E DISCUSSÃO

As Figuras 2 e 3 apresentam, no eixo Y, a vazão de dados dos QDR durante a Compensação de Movimento, expressa em GB/s, necessária para executar a decodificação à 30 qps. A Figura 1, no eixo X, faz a divisão inicial por Classes e a

secundária por QDR, enquanto separa cada imagem por QP. A Figura 2, no eixo X, faz a divisão inicial por Classes e a secundária por QP, enquanto separa cada imagem por QDR. A Classe A representa os vídeos de 3840 por 2160; a B, de 1920 por 1080; a C, de 1280 por 720; a D, de 832 por 480; e a E, de 416 por 240.

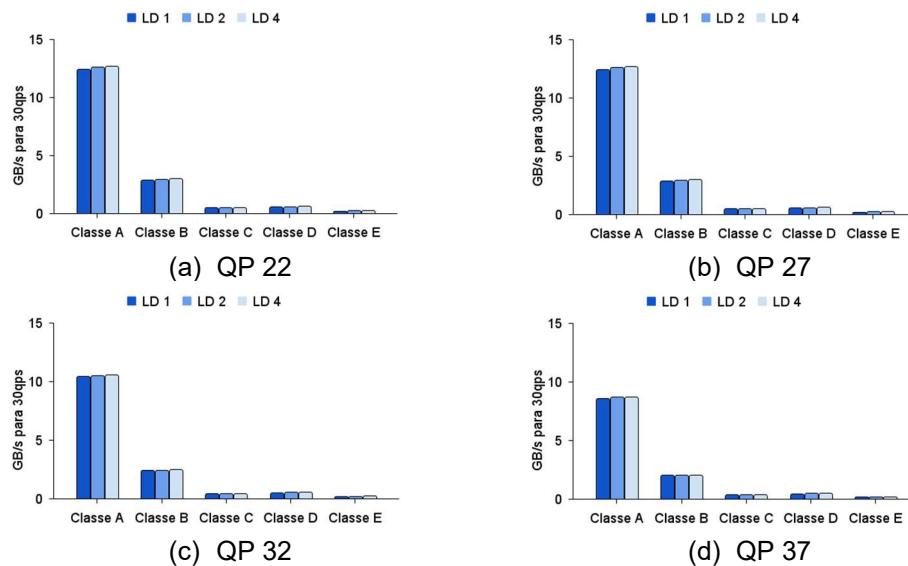
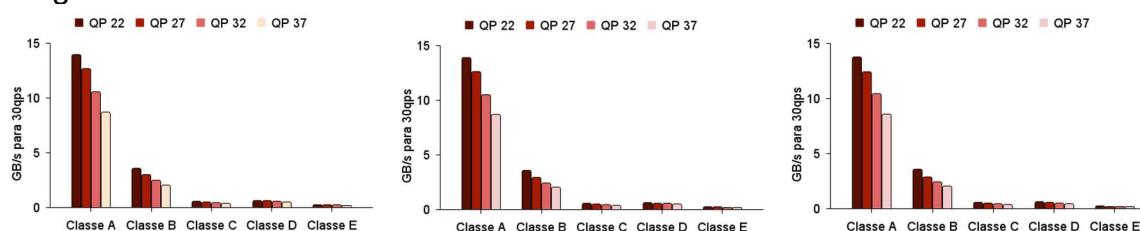


Figura 2: Taxa de transferência entre a memória e a Compensação de Movimento para 30 quadros por segundo.

De forma geral, maiores resoluções necessitam de taxas de transferências maiores, pois apresentam mais dados para serem processados dentro de cada quadro. A única exceção nos experimentos foi a Classe C, que teve sua média derrubada em cerca de 30% por vídeos relacionados a slides. Para um QP de 22 e LD 4 (padrão), as Classes de A à E apresentam uma necessidade média de 13,946, 3,531, 0,490, 0,597 e 0,207 GB/s, respectivamente, para atingir 30 qps em tempo real.

Em relação aos efeitos da redução dos números de QDR, a maioria apresentou uma redução nos acessos, com exceção, novamente, da Classe C, que apresentou um aumento de 1-2%. Entretanto, essas quedas foram, de forma geral, pouco expressivas. Mesmo em comparação com LD 1, que teve os melhores resultados, e QP 22, a média por classe não passou de uma redução de 3%, exceto pela Classe E, que viu uma redução de 11%. Além disso, conforme o QP aumenta, a redução de acessos tende a aumentar, mas a chance de vídeos terem um aumento no número de acessos também cresce.

A pequena redução de acessos com a limitação dos QDR provavelmente se deve a alta concentração dos acessos no quadro mais recentemente reconstruído. Pois, no Low Delay, ele é o mais temporalmente recente e, em geral, o mais similar, logo com mais e melhores redundâncias a serem aproveitadas durante a codificação. Enquanto os casos de aumentos de acessos, provavelmente, se devem a escolhas subótimas de candidatos durante a Estimação de Movimento, obrigando mais acessos à memória.



(a) LD 4

(b) LD 2

(c) LD 1

Figura 3: Taxa de transferência entre a memória e a Compensação de Movimento para 30 quadros por segundo.

Na questão dos efeitos dos QPs sobre os acessos, entre os de QP 22 e os de 27, 32 e 37, há uma queda em acessos de 11, 26 e 39%, respectivamente, em uma média de todas as classes. A quantidade de QDR tem um efeito pouco expressivo sobre essa variação, com o percentual flutuando dentro de 1%.

Essa alta variação de acessos com base no QP pode ser devida a um uso mais leve de técnicas de interpolação na Compensação de Movimento. Elas são usadas para gerar pixels intermediários entre dois já existentes, pois movimentos raramente acontecem com precisão em vídeos, mas sim de forma fracionada; e podem fazer a leitura de números diferentes de amostras de quadros de referência.

4. CONCLUSÕES

Este trabalho serve para motivar soluções que reduzam o consumo energético na comunicação de memória de um sistema de decodificação de vídeos de acordo com o padrão VVC, com foco nas memórias dos QDR. As restrições dos QDR tiveram efeitos insuficientemente expressivos para serem consideradas viáveis sozinhas como medida de redução do consumo energético. Entretanto, o aumento do QP se mostrou promissor e um potencial alvo de futuros estudos.

5. REFERÊNCIAS BIBLIOGRÁFICAS

YOUTUBE. Youtube Statistics. Youtube, San Bruno, out. 2018. Acessado em 24 jun. 2021. Online. Disponível em: <http://www.youtube.com/intl/pt-BR/about/press>

JVET. Versatile Video Coding (VVC) | JVET. Fraunhofer Heinrich Hertz Institute, Berlin, 2021. Acessado em 24 jun. 2021. Online. Disponível em: <http://jvet.hhi.fraunhofer.de>

ZHANG, F.; KATSENOU, A.V.; AFONSO, M.; DIMITROV, G.; BULL, D.R. Comparing VVC, HEVC and AV1 using objective and subjective assessments. **arXiv preprint arXiv:2003.10282**, 2020.

PAKDAMAN, F.; ADELIMANESH, M.A.; GABBOUJ, M.; HASHEMI, M.R. Complexity analysis of next-generation VVC encoding and decoding. **2020 IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP)**. IEEE, 2020. p. 3134-3138.

OHM. VTM common test conditions and software reference configurations for SDR video. WG 05 MPEG Joint Video Coding Team(s) with ITU-T SG 16, Virtual, 19 nov. 2020. Acessado em 12 mar. 2021. Online. Disponível em: https://jvet-experts.org/doc_end_user/current_document.php?id=10545

INTEL. Pin - A Dynamic Binary Instrumentation Tool. Intel Corporation, Santa Clara, 20 mai. 2021. Acessado em 06 ago. 2021. Online. Disponível em: <http://software.intel.com/content/www/us/en/develop/articles/pin-a-dynamic-binary-instrumentation-tool.html>