

AVALIAÇÃO DE PERFORMANCE DOS DEPTH MODELING MODES NA CODIFICAÇÃO EM MAPAS DE PROFUNDIDADE DO 3D-HEVC

CHRISTOPHER MOURA¹; MÁRIO SALDANHA¹; GUSTAVO SANCHEZ²; CÉSAR MARCON³; LUCIANO AGOSTINI¹; MARCELO PORTO¹

¹*Universidade Federal de Pelotas – ccmoura@inf.ufpel.edu.br, mrdfsaldanha@inf.ufpel.edu.br, agostini@inf.ufpel.edu.br, porto@inf.ufpel.edu.br*

²*IF Farroupilha – gustavo.sanchez@ifarroupilha.edu.br*

³*Pontifícia Universidade Católica do Rio Grande do Sul – cesar.marcon@pucrs.br*

1. INTRODUÇÃO

Diversas técnicas de compressão de vídeos 3D têm sido estudadas nos últimos anos. O *3D-High Efficiency Video Coding* (3D-HEVC) (TECH et al., 2016) é o codificador estado da arte para vídeos 3D, que foi desenvolvido pelo *Joint Collaborative Team on 3D Video Coding Extension Development* (JCT-3V) para alcançar maior eficiência em compressão de vídeos 3D que o padrão precedente.

A principal inovação do 3D-HEVC é o uso do *Multiview Video plus Depth* (MVD) (KAUF et al., 2007), onde cada *frame* de textura (imagem natural vista pelos usuários) tem um mapa de profundidade associado. Os mapas de profundidade são capturados por sensores infravermelhos acoplados na mesma câmera usada para captura das vistas de textura e são representados por imagens em escala de cinza, como demonstrado na Figura 1. Estes mapas de profundidade indicam a distância entre os objetos e a câmera. Mapas de profundidade são compostos por grandes regiões homogêneas (fundos e corpos de objetos) e arestas bem definidas (bordas dos objetos).



Figura 1 – (a) Frame de textura da sequência *Kendo* e (b) seu mapa de profundidade associado.

Com informações de textura e mapas de profundidades, técnicas leves de síntese de vistas tal como *Depth Image-based Rendering* (DIBR) (FEHN, 2004) podem ser usadas para sintetizar vistas intermediárias virtuais, onde as vistas virtuais são requeridas apenas no lado do decodificador, reduzindo o número de vistas de texturas transmitidas no vídeo codificado.

O 3D-HEVC é baseado no *High Efficiency Video Coding* (HEVC) (SULLIVAN et al., 2012) e herda muitas ferramentas de codificação, especialmente para vistas de textura. No entanto, novas ferramentas de codificação foram desenvolvidas para explorar redundâncias entre as vistas e executar a codificação em mapas de profundidade eficientemente.

Os *Depth Modeling Modes* (DMMs) (MERKLE et al., 2016) foram adicionados para codificar mapas de profundidade e gerar as vistas sintetizadas com alta qualidade. Os DMMs são compostos por dois novos modos, DMM-1 e DMM-4, que codificam regiões de arestas nos mapas de profundidade. Cada

modo divide o bloco de *pixels* em duas regiões que receberão um único valor chamado *Constant Partition Value* (CPV), calculado pela média das amostras de cada região. O DMM-1 divide o bloco por um *wedgelet* onde um *wedgelet* é uma linha reta que divide o bloco na região da aresta. O DMM-4 usa o bloco de textura colocalizado como referência para dividir o bloco. A Figura 2 ilustra estes dois modos.

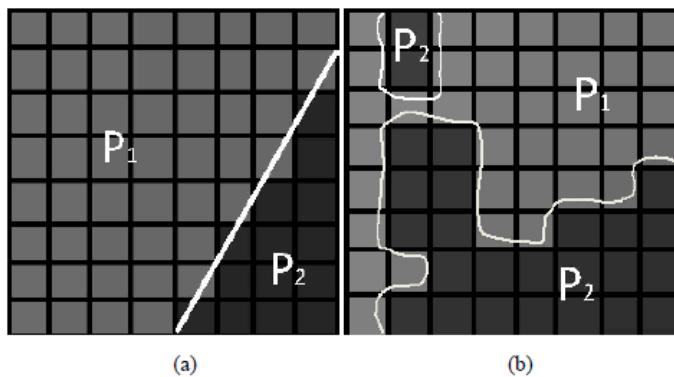


Figura 2 – (a) DMM-1 e (b) DMM-4.

Este trabalho tem como objetivo apresentar uma avaliação de performance considerando a eficiência e complexidade computacional (mensurado como tempo de execução) na codificação para mapas de profundidade no 3D-HEVC. Identificando o impacto dos DMMs, este trabalho contribuirá com outros trabalhos que focam em aumentar a eficiência de codificação ou reduzir a complexidade computacional de codificação dos mapas de profundidade.

2. METODOLOGIA

Para avaliar a performance dos DMMs no processo de codificação, foram feitas quatro simulações usando o *3D-HEVC Test Model* (3D-HTM) (versão 16.3) (3D-HEVC Test Model, 2016) conforme o *Common Test Conditions* (CTC) com a codificação All-intra do codificador. O CTC especifica oito sequências de vídeo para serem avaliadas, onde três vídeos têm um resolução de 1024x768 *pixels* e cinco vídeos têm uma resolução de 1920x1088 *pixels*. Cada sequência de vídeo foi codificada com quatro pares de *Quantization Parameter* (QP-texture, QP-depth) com valores (25, 34) (30, 39), (35, 42), e (40, 45).

As quatro simulações da análise foram: (i) Ambos DMMs ativados; (ii) apenas o DMM-1 desativado; (iii) apenas o DMM-4 desativado; e (iv) ambos DMMs desativados. A primeira simulação foi executada para obter os resultados do codificador 3D-HTM original. A segunda e a terceira simulação permite avaliar o impacto de cada DMM separadamente. A quarta simulação demonstra o impacto de usar ambos DMMs na codificação intra comparado com a primeira simulação.

A complexidade computacional foi mensurada usando o tempo de execução e a eficiência de codificação usando o *Bjontegaard Delta bitrate* (BDBR) (BJONTEGAARD, 2001) das vistas sintetizadas para indicar a variação na quantidade de bits para representar o vídeo codificado. Os resultados de cada avaliação foram coletados e comparados com os resultados de codificação sem nenhuma modificação, i.e., ambos DMMs ativados. Os experimentos foram executados em um servidor com o sistema operacional Ubuntu 16.04, memória de 94GB DDR3 e processador Intel® Xeon® E5-2650 com 48 cores e frequência de 2,20GHz.

3. RESULTADOS E DISCUSSÃO

A Figura 3 apresenta os resultados de eficiência. Quando o DMM-1 é desativado um alto impacto na codificação é ocasionado. Em média, o BDBR é aumentado em 3,0%. Desativando ambos DMMs o BDBR é aumentado em 5,6%. O maior impacto pode ser visto na sequência *Undo_Dancer* com um aumento no BDBR de 8,3%, enquanto que a sequência *Poznan_Street* obteve o menor impacto, com apenas 2,2% de aumento no BDBR.

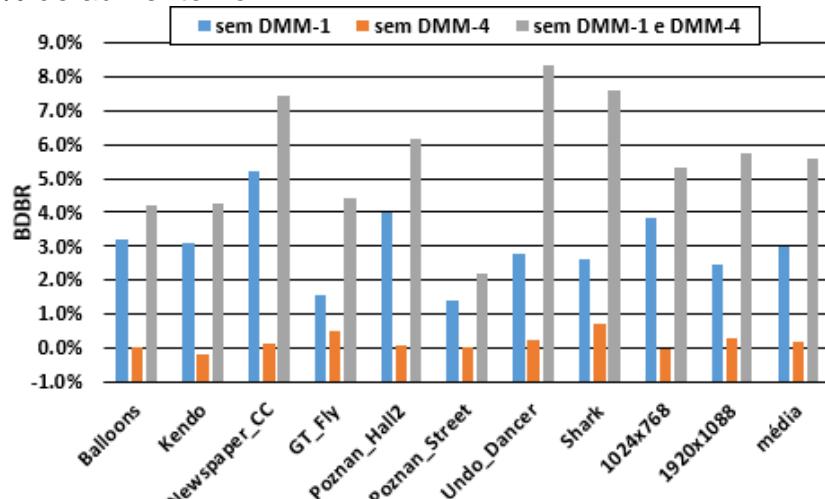


Figura 3 – Resultados em eficiência de codificação para ambos DMMs.

Os DMMs representam uma complexidade considerável na codificação. Quando desativamos qualquer um dos DMMs, uma redução no tempo de execução é esperada. Com o DMM-4 desativado, o tempo de execução foi reduzido em média em 4,7%. Quando desativamos o DMM-1, o tempo de execução é reduzido em média em 18,7%. Desativando ambos DMMs, a redução foi de 23,3.

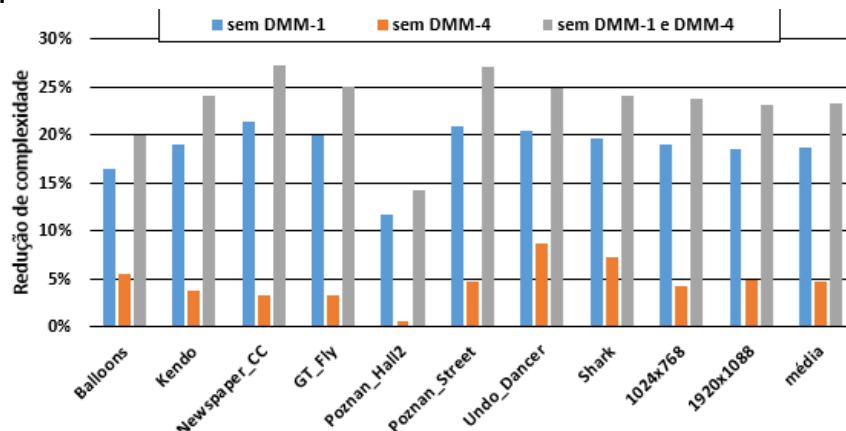


Figura 4 – Resultados em complexidade de codificação para ambos DMMs.

Na Figura 5, é mostrado que a redução de complexidade varia em diferentes cenários e que a redução tende a ser menor em valores altos de *QP-pair*. Para valores altos de *QP-pair*, o codificador comprime mais o vídeo, usando menos DMMs. Assim, a redução de complexidade é menor ao desativar os DMMs. Para valores baixos de *QP-pair*, o codificador usa mais DMMs, portanto a redução de complexidade é maior.

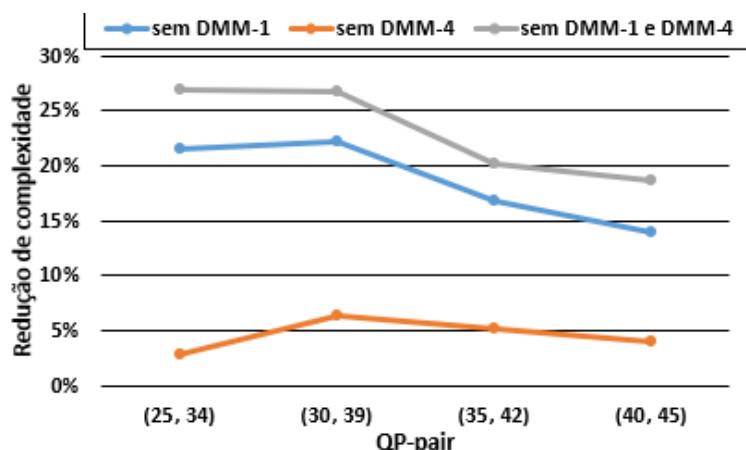


Figura 5 – Redução de complexidade por *QP-pair* ao desativar os DMMs.

4. CONCLUSÕES

Este trabalho apresentou uma avaliação de performance dos DMMs na codificação intra em mapas de profundidade do 3D-HEVC. A avaliação foi feita considerando a eficiência de codificação e complexidade computacional para os casos definidos no CTC. A avaliação apresentou resultados que podem ser usados como referência em soluções focadas na redução de complexidade em mapas de profundidade para o codificador 3D-HEVC.

5. REFERÊNCIAS BIBLIOGRÁFICAS

TECH, G. et al. Overview of the Multiview and 3D Extensions of High Efficiency Video Coding. **Ieee Transactions On Circuits And Systems For Video Technology**, [S.L.], v. 26, n. 1, p. 35-49, jan. 2016.

KAUFF, P. et al. Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability. **Signal Processing: Image Communication**, [S.L.], v. 22, n. 2, p. 217-234, fev. 2007.

FEHN, C. Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV. **Stereoscopic Displays And Virtual Reality Systems XI**, [S.L.], v. 5291, p. 93-104, 21 maio 2004.

SULLIVAN, G. et al. Overview of the High Efficiency Video Coding (HEVC) Standard. **Ieee Transactions On Circuits And Systems For Video Technology**, [S.L.], v. 22, n. 12, p. 1649-1668, dez. 2012.

MERKLE, P. et al. Depth Intra Coding for 3D Video Based on Geometric Primitives. **Ieee Transactions On Circuits And Systems For Video Technology**, [S.L.], v. 26, n. 3, p. 570-582, mar. 2016.

3D-HEVC Test Model, JCT-VC. 21 mar. 2016. Acessado em 14 set. 2020. Online. Disponível em: https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-16.3/.

BJONTEGAARD, G. Calculation of Average PSNR Differences between RD-Curves. **Proceedings of the ITU-T Video Coding Experts Group (VCEG) Thirteenth Meeting**, [S.L.], p. 1-5, 2001.