

ANÁLISE DA EFICIÊNCIA DOS MODOS DMMS NA PREDIÇÃO INTRA- QUADROS DE MAPAS DE PROFUNDIDADE NO 3D-HEVC

GABRIEL BALOTA; MÁRIO SALDANHA; GUSTAVO SANCHEZ ;MARCELO
PORTO; LUCIANO AGOSTINI; BRUNO ZATT

Universidade Federal de Pelotas
{mrdfsaldanha, gmbalota, gfsanchez, agostini, porto, zatt}@inf.ufpel.edu.br
Grupo de Arquiteturas e Circuitos Integrados

1. INTRODUÇÃO

Nos dias atuais, cada vez mais cresce a diversidade de dispositivos que suportam aplicações de vídeo em alta definição (HD), como por exemplo, vídeos 3D. A codificação desses vídeos 3D é essencial, pois eles possuem uma elevada taxa de dados que necessitam ser transmitidas e armazenadas. Por exemplo, em um vídeo 3D com duração de 20 minutos, e resolução HD 1080p, a uma taxa de exibição 30 quadros por segundo, com cinco vistas de textura sem compressão, são necessários 350 GB para armazená-lo. Com a adição de canais de profundidade, tipo de dado utilizado no padrão emergente *3D-High Efficiency Video Coding* (3D-HEVC) (TECH, 2013), o volume de dados cresce em 33,3%. Com este grande volume de dados a complexidade computacional para codificar um vídeo é muito alta. Neste contexto, o desenvolvimento de heurísticas para reduzir a complexidade da codificação de vídeo 3D se torna interessante.

O padrão emergente 3D-HEVC é uma extensão para o padrão de codificação de vídeo 2D, o *High Efficiency Video Coding* (HEVC) (ITU-T, 2013). Algoritmos para codificação de vídeos 2D não possuem a característica de explorar as redundâncias entre vista no mesmo instante de tempo. Com isso novas ferramentas foram adicionadas para atingir maior eficiência na codificação em termos de compressão e qualidade.

Algumas ferramentas inseridas no novo codificador 3D são a estimação de disparidade, loop de filtragem pós-processamento, predição inter-vistas baseada em síntese de vista e compensação de iluminação, entre outras.

Os mapas de profundidade são informações adicionais, enviadas em conjunto com as informações de textura (vídeo convencional). No padrão emergente 3D-HEVC, cada quadro de textura possui um mapa de profundidade associado, e devido a isso, a codificação dos mapas de profundidade foi inserida no padrão. Os mapas de profundidade são compostos por imagens em tons de cinza com 8 bits por amostra, que tem como objetivo informar a distância entre os objetos e a câmera. São caracterizados por possuírem regiões bastante homogêneas e suas arestas são bem definidas como podemos ver na Figura 1. Os valores dos bits variam do valor de 0 a 255.

A vantagem de utilizar os mapas de profundidade na codificação de vídeos 3D é a possibilidade de gerar, de forma eficiente, vistas intermediárias que não precisam ser transmitidas e nem armazenadas através da interpolação dos canais de textura com os canais de profundidade. (MERKLE, 2011)

Por um lado, a predição intra-quadro herdada do HEVC codifica de forma eficiente áreas homogêneas, gerando ótimos resultados. Porém as arestas são codificadas de forma ineficiente, geram artefatos (ruído) nas vistas intermediárias (MULLER, 2013). Corroborando este fato, os desenvolvedores do padrão desenvolveram novos modos para a predição intra-quadro de mapas de profundidade: os modos *Depth Modeling Modes* (DMM). Nos DMMs, o bloco de profundidade é subdividido em duas regiões, onde cada região é representada

por um valor constante. Esses novos modos são utilizados somente para a codificação de mapas de profundidade.

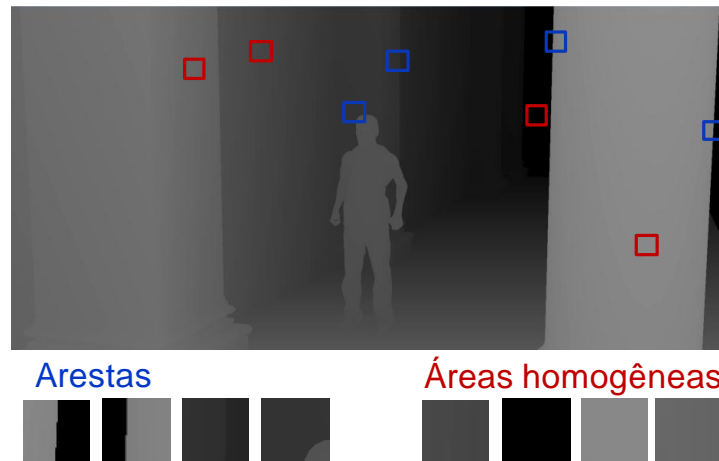


Figura 1 – Mapa de profundidade da sequência Undo_Dancer.

Este trabalho tem o objetivo de apresentar uma avaliação quanto à relevância da utilização dos modos DMM na predição intra-quadros dos mapas de profundidade na codificação de um vídeo 3D, usando o padrão 3D-HEVC.

2. METODOLOGIA

A codificação dos mapas de profundidade representa cerca de 24,3% de toda complexidade na codificação de um vídeo 3D, como é possível visualizar na Figura 2. Com o objetivo de entender a eficiência de cada um dos modos DMM em termo de qualidade de vídeo, taxa de compressão e complexidade computacional, foram avaliados os modos DMM 1 e DMM 4 (que são os únicos modos que se encontram disponíveis na atual versão do software de referência).

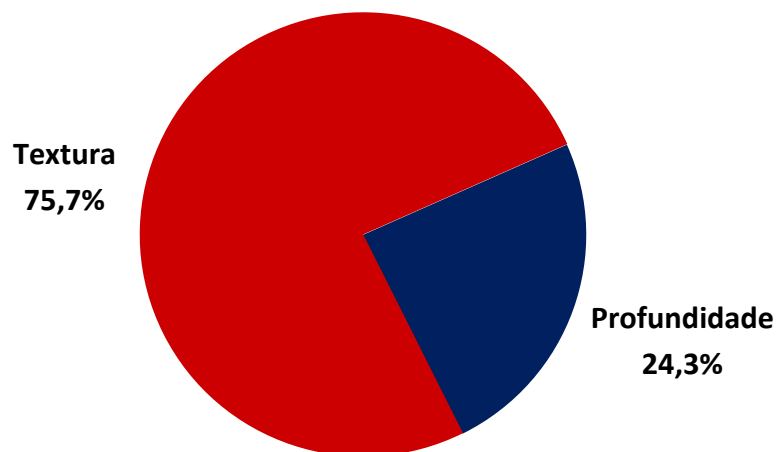


Figura 2 – Complexidade da codificação.

Uma das estratégias utilizadas para executar a divisão do bloco de profundidade é através de uma partição *wedgelet*, DMM 1, apresentado na Figura 3(a). Na partição *Wedgelet* as regiões são separadas por uma reta, formando duas regiões, P_1 e P_2 . Nota-se que a reta corta algumas amostras no meio, porém cada amostra é mapeada por um valor binário sinalizando a qual região a mesma pertence.

Outra estratégia é utilizar contorno para fazer a divisão das regiões, DMM 4, como mostrado na Figura 3(b). Utilizando esta estratégia as regiões podem ter formatos aleatórios e podem ser compostos de várias partes. Para gerar o contorno é utilizado como referência o quadro de textura co-localizado.

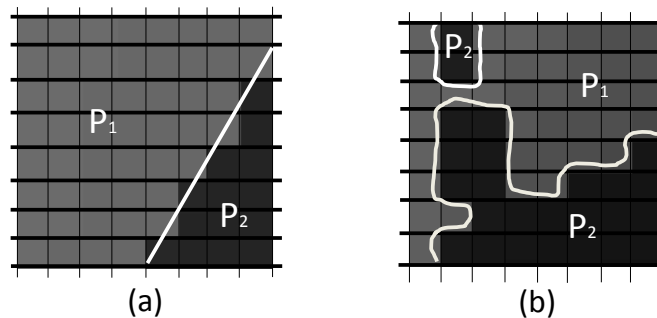


Figura 3 – (a) Subdivisão por wedgelet (b) Subdivisão por contorno.

3. RESULTADOS E DISCUSSÃO

As avaliações foram geradas a partir do software de referência do padrão emergente o 3D-HEVC, o 3D-HEVC Test Model (HTM) versão 10.2, utilizando as condições comuns de teste (CTC) (TECH, 2013). A Tabela 1 apresenta a partir dos resultados obtidos pelo software de referência, uma comparação do BD-Rate entre os testes executados. A comparação do BD-Rate é feita com a comparação da qualidade para o mesmo *bitrate* dos vídeos.

Tabela 1 – Avaliação de BD-rate do canal de profundidade.

Vídeos	Sem DMM	Sem DMM 1	Sem DMM 4
Balloons	1,4%	1,2%	0,1%
Kendo	1,4%	1,1%	0,2%
Newspaper_CC	4,0%	3,1%	0,4%
GT_Fly	0,8%	0,8%	-0,1%
Poznan_Hall2	2,1%	1,6%	0,1%
Poznan_Street	1,3%	1,0%	0,1%
Undo_Dancer	4,2%	2,0%	0,5%
Média	2,1%	1,5%	0,2%

Analisando a Tabela 1 pode-se perceber que não nenhum ganho na retirada dos modos DMM, e sim perdas na qualidade dos vídeos. Em média obteve-se perda de 2,1% na qualidade. Também é possível perceber que ao tirar somente o DMM4 do codificador obteve-se uma perda pequena, e até mesmo um ganho no vídeo GT_Fly, em relação aos outros testes realizados.

Como citado anteriormente, esses testes têm o objetivo medir a eficiência de cada modo DMM na codificação dos mapas de profundidade. A Tabela 2 apresenta dados da redução de complexidade com a retirada dos modos DMM.

Com a Tabela 2 é possível notar que todos os testes executados tiveram o objetivo alcançado. Com 12,6% da redução da complexidade da codificação dos mapas de profundidade, a retirada dos dois modos DMM, foi o que obteve melhor resultado em redução de complexidade, porém foi o teste que obteve o pior

resultado de qualidade. Já a retirada do modo 4 que alcançou o pior desempenho em redução de qualidade, foi o que melhor alcançou os resultados de qualidade.

Tabela 2 – Média da redução da complexidade do canal de profundidade

Vídeos	Sem DMM	Sem DMM 1	Sem DMM 4
Balloons	12,2%	8,4%	5,4%
Kendo	12,4%	9,8%	5,9%
Newspaper_CC	14,6%	11,7%	6,7%
GT_Fly	11,9%	9,8%	4,2%
Poznan_Hall2	13,1%	10,2%	5,5%
Poznan_Street	12,8%	9,9%	5,4%
Undo_Dancer	11,0%	8,4%	4,8%
Média	12,6%	9,7%	5,4%

4. CONCLUSÕES

Os modos DMM foram criados com o objetivo de melhorar a predição intra-quadros, pois os modos herdados no HEVC são bastante eficientes para áreas homogêneas, porém, podem gerar artefatos nas vista intermediárias ao codificarem áreas que possuem arestas. Grande parte da complexidade do codificador 3D está na codificação dos mapas de profundidade, onde atuam os modos DMM.

Este trabalho apresentou um estudo da eficiência de cada modo DMM na codificação dos vídeos 3D, cortando a execução dos modos DMM. Com o corte dos modos foi possível obter o objetivo principal, a redução de complexidade, porém não foi possível reduzi-la mantendo a mesma qualidade.

Para trabalhos futuros pretende-se criar estratégias com o objetivo de reduzir a complexidade da codificação dos mapas de profundidade, explorando, além dos modos DMM, outros modos da codificação sem perdas de qualidade.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- TECH, G.; WEGNER, K.; CHEN, Y.; YEA, S. **3D HEVC Test Model 3**. Documento: JCT3V-C1005. Draft 3 of 3D-HEVC Test Model Description. Geneva, 2013.
- ITU-T. **SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS. Infrastructure of audiovisual services – Coding of moving video**. High efficiency video coding 2013
- MERKLE, P.; et al. **3D video: Depth Coding Based on Inter-Component Prediction of Block Partitions**. IEEE Picture Coding Symposium (PCS), 2011.
- MULLER, K. et al. **3D High-Efficiency Video Coding for Multi-View Video and Depth Data**. IEEE Transactions on Image Processing, V. 22, n. 9, p. 3366 – 3378, 2013.