

REDUÇÃO DA COMPLEXIDADE DA CODIFICAÇÃO DE MAPAS DE PROFUNDIDADE NO PADRÃO EMERGENTE 3D-HEVC

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

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

1. INTRODUÇÃO

As aplicações 3D estão se tornando mais populares nos últimos anos. Isto está diretamente relacionado ao fato de que o custo dos dispositivos que armazenam e/ou reproduzem vídeos 3D está diminuindo e o interesse na pesquisa e no desenvolvimento de tecnologias para codificar vídeos 3D está aumentando.

Os vídeos 3D são compostos de múltiplas vistas e geram uma grande quantidade de dados, principalmente em vídeos de alta definição (*High Definition - HD*), que necessitam ser armazenados ou transmitidos. Com base neste fato, é necessário o estudo e o desenvolvimento de algoritmos eficientes para codificação destes vídeos.

A codificação de mapas de profundidade é uma nova ferramenta introduzida no padrão emergente 3D-*High Efficiency Video Coding* (3D-HEVC) (TECH, 2013), utilizada na codificação de vídeos 3D, onde cada quadro de textura está associado a um mapa de profundidade. Os mapas de profundidade são representados de forma monocromática, ou seja, utiliza-se tons de cinza com 8 bits para representação, que variam do valor 0 até o valor 255, onde pontos pertos da câmera possuem seus valores mais próximos de 255, enquanto pontos mais distantes possuem seus valores próximos de 0. A Figura 1 mostra um exemplo de um mapa de profundidade retirado da sequência *Poznan_Street*. É possível notar que objetos mais distantes da câmera são representados por tons mais escuros, enquanto objetos mais próximos da câmera são representados por tons mais claros.



Figura 1 – Imagem de profundidade do vídeo *Poznan_Street*

A vantagem da utilização de mapas de profundidade associados a quadros de textura é que é possível gerar, eficientemente, vídeos de pontos de vista intermediários através da interpolação de canais de textura e de profundidade (MERKLE, 2007).

Nos mapas de profundidade é possível identificar facilmente os objetos presentes na imagem. Isto é uma característica importante dos mapas de profundidade, eles possuem uma grande quantidade de regiões homogêneas e arestas bem definidas.

A predição intra-quadro tradicional, empregada nos padrões de codificação de vídeo 2D, obtém bons resultados codificando as áreas homogêneas, porém a codificação de arestas utilizando a predição intra-quadro tradicional pode resultar em artefatos nas vistas intermediárias sintetizadas (MULLER, 2011). Com isto, o padrão emergente 3D-HEVC define quatro novos tipos de predição intra-quadro para codificação de profundidade, os modos de modelagem de profundidade (DMM – *Depth Modeling Modes*). Nos quatro modos, o bloco de profundidade é subdividido em duas regiões, onde cada região é representada por um valor constante.

Uma das formas de fazer esta divisão é através de uma partição *Wedgelet*, apresentada na Figura 2. Na partição *Wedgelet* as duas regiões são separadas por uma linha reta que separa as regiões P_1 e P_2 . É possível notar que a reta passa no meio de algumas amostras e cada amostra é mapeada para um valor binário que identifica a qual região ela pertence.

Entre os quatro modos de codificação de mapas de profundidade o foco deste artigo é o modo um, onde a escolha da reta é feita exaustivamente em uma busca sobre um conjunto de partições *Wedgelet*. A partição *Wedgelet* que alcançar a menor distorção (erro de codificação) entre o sinal original e a predição através da *Wedgelet* é selecionada. Com base neste fato, um algoritmo de busca eficiente é importante para diminuir a complexidade do codificador.

Este trabalho apresenta o *Gradient Simplified Mode One* (GSMO), que é a aplicação de uma heurística para simplificar esta etapa citada anteriormente, com objetivo de minimizar a complexidade da seleção da partição *Wedgelet* de acordo com o tamanho do bloco.

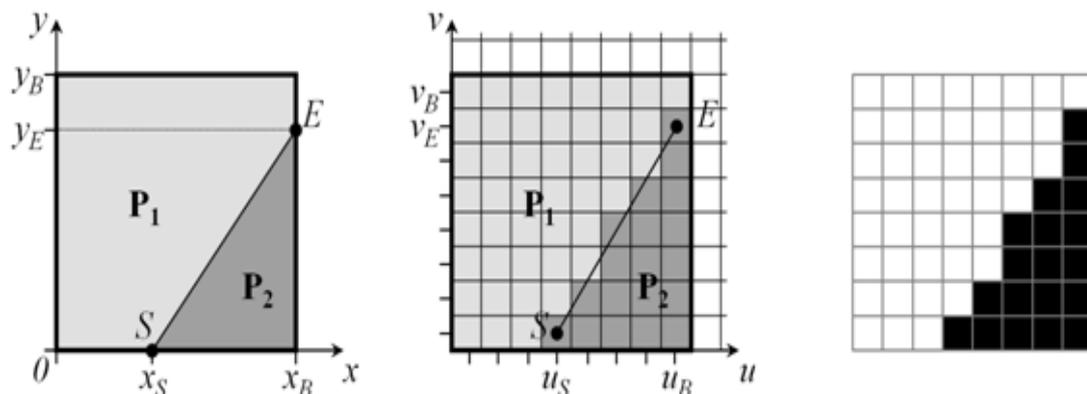


Figura 2 – Predição de profundidade por *Wedgelet* (MERKLE, 2011).

2. METODOLOGIA

A Tabela 1 apresenta todas as possibilidades de partições *Wedgelet* e as partições que são avaliadas para cada tamanho de bloco (as partições possíveis englobam as avaliadas e os refinamentos). Através desta análise é possível notar que são avaliadas muitas possibilidades de partições *Wedgelet*. Além disso, motivado pelo fato de que as partições *Wedgelet* foram criadas para separar duas regiões com um gradiente abrupto foi desenvolvida uma simplificação no conjunto de escolha das partições *Wedgelet*, onde foi aplicada uma heurística que decide o ponto inicial de onde as retas que definem as partições *Wedgelet* devem iniciar.

Primeiramente é feita uma comparação entre todos os pares de pixels vizinhos das bordas, do bloco de profundidade avaliado. O objetivo desta etapa é encontrar o ponto da borda do bloco em que existe o maior gradiente entre os pixels vizinhos. Com isso, é possível dizer que este ponto possui uma alta probabilidade de ser a origem da reta *Wedgelet* utilizada. Após obter a coordenada são avaliadas todas as possibilidades de retas partindo do ponto selecionado, e por fim é selecionada a partição *Wedgelet* dentro deste subconjunto que obtém a menor distorção entre o sinal original e a aproximação *Wedgelet* avaliada.

Tabela 1 – Número de partições *Wedgelet* para cada tamanho de bloco.

Tamanho do bloco	Número de partições <i>Wedgelet</i> possíveis	Número de partições <i>Wedgelet</i> avaliadas
4x4	86	58
8x8	782	310
16x16	1394	338
32x32	1503	368

3. RESULTADOS E DISCUSSÃO

As avaliações foram feitas utilizando o software de referência 3D-HEVC *Test Model* (HTM) e as condições comuns de testes (CTC) (TECH, 2013) foram utilizadas para avaliar o desempenho do algoritmo proposto em relação aos algoritmos tradicionais propostos pelo padrão. Os resultados são apresentados na Tabela 2. Os primeiros três vídeos possuem resolução 1024x768 pixels e os próximos quatro vídeos são *High Definition* (HD) 1080p.

Tabela 2 – Avaliação de BD-rate, tempo de codificação e complexidade do canal de profundidade

Vídeos	PSNR textura / bitrate total	PSNR vista sintetizada / bitrate total	Redução do Tempo Total de codificação	Redução da complexidade de profundidade
Balloons	-0,1%	-0,4%	2,4%	9,5%
Kendo	0,0%	-0,2%	2,7%	9,4%
Newspaper_CC	0,2%	-0,3%	2,5%	9,6%
GT_Fly	0,0%	-0,2%	2,7%	11,0%
Poznan_Hall2	0,0%	-0,3%	2,6%	12,3%
Poznan_Street	0,1%	-0,2%	3,2%	11,9%
Undo_Dancer	-0,1%	-0,4%	2,7%	12,3%
Média	0,0%	-0,3%	2,7%	10,9%

Os canais de profundidade representam em média 24,30% da complexidade computacional de todo o codificador. Analisando a Tabela 2 é possível notar que a simplificação que foi aplicada nos mapas de profundidade diminuiu em média 10,9% da complexidade dos canais de profundidade. Além disso, foi possível reduzir em média 2,7% do tempo total de toda a codificação, alcançando uma perda aceitável de bitrate em média de 0,3% para uma mesma qualidade de vídeo nas vistas sintetizadas.

Nas sequências com resolução de 1024x768 pixels foi possível alcançar uma redução da complexidade de profundidade e no tempo total de codificação de em média 9,5% e 2,5%, respectivamente, ocorrendo uma perda de bitrate de apenas 0,3% para uma mesma qualidade de vídeo.

Analisando os resultados gerados para as sequências HD 1080p é possível notar um ganho ainda maior na redução da complexidade dos canais de profundidade de em média 11,9% e também na redução do tempo de codificação em média de 2,8%, obtendo a mesma perda de 0,3% de bitrate para uma mesma qualidade vídeo.

Além disso, é possível observar que os maiores ganhos ocorreram nos vídeos HD 1080p. As sequências Ponzan_Hall2 e Undo_Dancer alcançaram uma redução na complexidade dos canais de profundidade de 12,3%, e a sequência Poznan_Street obteve uma redução do tempo de codificação de 3,2%.

4. CONCLUSÕES

A predição intra-quadro tradicional se torna eficiente em áreas homogêneas e permite alta eficiência de codificação nessas áreas com valores quase constantes, contudo em áreas que possuem arestas, seu uso pode gerar artefatos nas vistas intermediárias sintetizadas. Baseado neste fato, o padrão 3D-HEVC definiu novos modos de predição intra-quadro para codificação dos mapas de profundidade.

Este trabalho apresentou uma simplificação no algoritmo de um dos modos de predição intra-quadro na codificação dos mapas de profundidade. Onde foi possível obter uma redução da complexidade de em média 10,9% na codificação dos canais de profundidade. Ao aplicar o algoritmo proposto, a perda no bitrate para uma mesma qualidade foi de 0,3%.

Como trabalhos futuros pretende-se aperfeiçoar esta simplificação, sem realizar uma simplificação tão rígida. Além disso, pretende-se explorar também os outros modos de codificação dos mapas de profundidade, com o objetivo de obter resultados ainda mais significativos.

5. REFERÊNCIAS BIBLIOGRÁFICAS

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.

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.