

ANÁLISE DA PREDIÇÃO INTER-QUADROS DOS MAPAS DE PROFUNDIDADE NO 3D-HEVC

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

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

1. INTRODUÇÃO

Nos últimos anos ocorreu uma popularização no cenário de vídeos 3D, principalmente com televisores e cinemas que suportam este tipo de vídeo. No entanto, estes vídeos geram uma grande quantidade de dados, quando comparados a vídeos 2D, devido as múltiplas câmeras necessárias para capturar o vídeo. Este volume de dados aumenta ainda mais na codificação de vídeos de alta definição. Com esta popularização e o desafio de manipular uma grande quantidade de dados, surge o interesse em pesquisar e desenvolver algoritmos capazes de codificar os vídeos 3D de forma eficiente.

Explorando os algoritmos de codificação mais promissores um padrão de vídeos 3D está em desenvolvimento, conhecido como padrão emergente 3D-*High Efficiency Video Coding* (3D-HEVC) (TECH, 2013). O 3D-HEVC é uma extensão do padrão bastante conhecido de codificação de vídeos 2D, *High Efficiency Video Coding* (HEVC), onde são adicionadas novas ferramentas para suportar a codificação de vídeos 3D.

Os vídeos 3D são compostos por múltiplas vistas, de múltiplas câmeras, criando a necessidade de codificar um conjunto de vídeos simultaneamente. Este processo acaba ocasionando em um grande aumento na complexidade de codificação e na largura de banda necessária para transmissão de um vídeo 3D, quando comparado aos vídeos 2D.

Para reduzir a largura de banda necessária para transmissão dos vídeos 3D, o padrão emergente 3D-HEVC adota o formato de dados *Multi-View plus Depth* (MVD). No formato MVD cada quadro do vídeo (conhecido como quadro de textura) está associado a um mapa de profundidade. Assim, somente um subconjunto das câmeras, juntamente com seus mapas de profundidade, presentes na cena são utilizadas para codificação, dispensando a necessidade de codificar todas as câmeras. Com as informações de textura e profundidade é possível gerar, de forma eficiente, vídeos de pontos de vista intermediários através da interpolação dos canais de textura e de profundidade (MERKLE, 2007).

O mapas de profundidade são representados utilizando amostras de 8 bits e tipicamente são visualizados como imagens em tons de cinza, representando a distância dos objetos e a câmera. Objetos distantes da câmera são representados por tons de cinza mais escuro, enquanto objetos próximos da câmera são representados por tons de cinza mais claro. É importante destacar que os mapas de profundidade possuem duas características bem definidas: regiões bastante homogêneas (frequentemente dentro de objetos) e arestas bem definidas (frequentemente na borda de objetos).

Como os algoritmos do HEVC foram desenvolvidos para manipular dados de textura, estes mesmos algoritmos quando utilizados para codificação dos mapas de profundidade podem apresentar pouca eficiência ou até mesmo um aumento

na complexidade, devido as informações dos mapas de profundidade serem mais simples, quando comparado aos dados de textura.

A predição inter-quadros realizada nos mapas de profundidade é semelhante a predição realizada com dados de textura, aplicando a estimação de movimento (*Motion Estimation* - ME) com o objetivo de encontrar o melhor casamento de bloco para cada bloco presente no quadro que está sendo codificado, buscando em quadros de referência. Para avaliar o melhor casamento de bloco é utilizado um critério de similaridade, onde o *Sum of Absolute Differences* (SAD) é um dos critérios mais utilizados.

Um algoritmo de busca é necessário para encontrar o casamento dos blocos, no software de referência do 3D-HEVC, 3D-HEVC *Test Model* (HTM), o algoritmo aplicado é o *TZ Search* (TZS) (TANG, 2010) tanto para as informações de textura, quanto para as informações de profundidade. Entretanto, considerando a simplicidade dos dados dos mapas de profundidade é possível pressupor que algoritmos mais simples que o TZS podem ser aplicados para reduzir a complexidade, mantendo uma boa eficiência na codificação. Portanto, este trabalho tem como objetivo apresentar uma avaliação realizada na predição inter-quadros dos mapas de profundidade aplicando algoritmos conhecidos na literatura, como: *Diamond Search* (DS) (ZHU, 2000) e *Hexagon Search* (HS) (ZHU, 2002).

2. METODOLOGIA

A Figura 1 (a) e (b) (SALDANHA, 2015) apresenta os mapas de calor para o SAD na estimação de movimento de um bloco de textura e um de mapa de profundidade, respectivamente. Regiões com azul escuro representam a melhor predição, ou seja valores baixos de SAD, enquanto regiões com cor vermelha representam os piores casamentos de bloco. Analisando a Figura 1 (a) é possível notar que o mapa de calor contendo informações de textura apresenta um comportamento complexo, com variações abruptas e múltiplos vales e picos. Estas características acabam ocasionando o problema dos mínimos locais (SANCHEZ, 2012). Baseado neste fato, algoritmos rápidos para a predição inter-quadros de textura, como *Diamond Search* (DS) e *Hexagon Search* (HS), podem não alcançar uma boa eficiência devido aos problemas dos mínimos locais (SANCHEZ, 2012). Com isso, existe a necessidade de utilizar-se algoritmos mais complexos, como por exemplo, o algoritmo TZS.

No entanto, analisando o mapa de calor para mapa de profundidade associado na Figura 1 (b) é possível notar regiões com valores mais semelhantes, e com variações mais suaves. Neste caso, algoritmos rápidos são capazes de encontrar um valor ótimo ou próximo do ótimo para o SAD sem realizar muitas iterações.

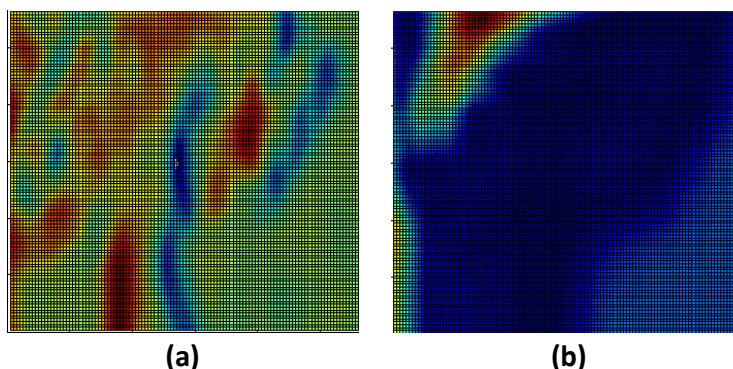


Figura 1 – Mapa de calor (a) de textura (b) de profundidade (SALDANHA, 2015).

Com esta análise é possível notar que os dados de textura apresentam um comportamento complexo e necessitam de algoritmos de busca mais sofisticados para evitar o mínimo local, porém, para os mapas de profundidade padrões simples devem ser o suficiente para lidar com suas características.

No software de referência do 3D-HEVC (HTM) é utilizado o algoritmo TZS tanto para textura quanto para os mapas de profundidade. Considerando os fatos citados, um algoritmo mais simples que o TZS pode ser aplicado para reduzir a complexidade sem apresentar grandes perdas na eficiência da codificação.

Com base nesta análise foi realizada a implementação e a avaliação dos algoritmos DS e HS, no software de referência 3D-HEVC *Test Model* (HTM), versão 10.2. Nos experimentos foram utilizadas as condições comuns de teste (*Common Test Conditions* - CTC) dos vídeos 3D.

3. RESULTADOS E DISCUSSÃO

A Tabela 1 apresenta os resultados em percentuais das vistas sintetizadas (BD-rate), tempo de codificação e a redução do SAD, que representa a redução em número de blocos candidatos avaliados (utilizando o critério de similaridade SAD), para os algoritmos DS e HS, quando comparados ao HTM sem modificações, ou seja, com a utilização do algoritmo TZS para a predição interquadros.

Tabela 1 – Resultados da avaliação dos algoritmos rápidos na ME.

Vídeos	Diamond Search (DS)			Hexagon Search (HS)		
	Vistas Sintetizadas (BD-rate)	Tempo de Codificação	Redução SAD	Vistas Sintetizadas (BD-rate)	Tempo de Codificação	Redução SAD
Balloons	0,028%	92,08%	84,75%	0,064%	92,10%	90,45%
Kendo	0,171%	91,16%	87,56%	0,094%	91,15%	92,07%
Newspaper_CC	0,011%	91,83%	87,40%	0,113%	91,89%	91,71%
GT_Fly	0,403%	92,93%	82,03%	0,406%	92,93%	88,95%
Poznan_Hall2	0,345%	92,65%	80,50%	0,436%	92,54%	87,08%
Poznan_Street	0,048%	92,86%	82,32%	0,091%	92,86%	88,83%
Undo_Dancer	0,431%	93,65%	78,28%	0,512%	93,51%	85,96%
1024x768	0,070%	91,69%	86,57%	0,090%	91,70%	91,41%
1920x1088	0,307%	93,02%	80,78%	0,361%	92,96%	87,70%
Média	0,205%	92,45%	83,26%	0,245%	92,42%	89,29%

Analisando a Tabela 1 é possível notar que o algoritmo DS aumenta o BD-rate nas vistas sintetizadas em 0,205%, em média. Este impacto é aceitável, considerando que o DS foi capaz de fornecer uma considerável redução na complexidade de 7,55%, em média, considerando todo codificador. Além disso, o algoritmo DS é capaz de reduzir os cálculos de SAD em 83,26%.

Considerando o algoritmo HS, em média ocorreu um aumento no BD-rate das vistas sintetizadas de 0,245%. Este aumento é aceitável, visto que o HS pode fornecer uma redução no tempo de codificação de 7,58%, em média, considerando os canais de profundidade e textura. Além disso, também como o DS, o HS é capaz de reduzir consideravelmente os cálculos do SAD, em 89,29% em média.

Com os resultados apresentados é possível concluir que os mapas de profundidade possuem padrões simples e que algoritmos que apresentam menor esforço computacional podem ser utilizados para reduzir o tempo de codificação com um impacto mínimo na qualidade do vídeo codificado. Além disso, esta expressiva redução nos cálculos do SAD, em ambos algoritmos, beneficia o desenvolvimento em *hardware*, pois reduz o número de acessos à memória para processar os blocos e então realizar o cálculo dos valores do SAD.

4. CONCLUSÕES

Este trabalho teve como objetivo analisar as características dos mapas de profundidade presentes no padrão emergente 3D-HEVC, com foco na estimação de movimento. Através da análise realizada foi possível notar que algoritmos mais simples do que os utilizados para codificar dados de textura podem alcançar uma grande eficiência na codificação dos mapas de profundidade.

Os algoritmos DS e HS foram implementados e avaliados no HTM, com esta avaliação foi possível concluir que estes algoritmos são capazes de reduzir consideravelmente o tempo de codificação com um impacto mínimo na qualidade do vídeo codificado. Além disso, reduzem consideravelmente o número de blocos comparados, facilitando o projeto em *hardware*, reduzindo o número necessário de acessos à memória.

Como trabalhos futuros, pretende-se aplicar esta análise para elaboração de um algoritmo para a predição inter-quadros dos mapas de profundidade, com a finalidade de aplicar padrões ainda mais simples do que os apresentados neste trabalho, com o objetivo de reduzir a complexidade computacional com um impacto aceitável na qualidade do vídeo codificado.

5. REFERÊNCIAS BIBLIOGRÁFICAS

SALDANHA, M.; SANCHEZ, G.; ZATT, B.; PORTO, M.; AGOSTINI, L.. **Complexity Reduction for the Emerging 3D-HEVC Depth Maps Coding**. In: IEEE International Symposium on Circuits and Systems, 2015, Lisboa. Proc. of ISCAS 2015, 2015.

Sanchez, G.; et al., **DMPDS: A Fast Motion Estimation Algorithm Targeting High Resolution Videos and Its FPGA Implementation**, IJRC, v. 2012, p. 1-12, 2012.

TECH, G.; et. al. **3D HEVC Test Model 3. Documento: JCT3V-C1005**. Draft 3 of 3D-HEVC Test Model Description. Geneva, 2013.

X.L. Tang, S.K. Dai, C.H. Cai, **An analysis of TZSearch Algorithm in JMVC**, ICGCS, pp. 516-520, 2010.

Zhu, S.; Ma, K. 2000. **A New Diamond Search Algorithm for FastBlock-Matching Motion Estimation**, IEEE TIP, vol. 9, n. 2, pp. 287- 290, Feb 2000.

Zhu, C.; Lin, X.; Chau, L., **Hexagon-Based Search Pattern for Fast Block Motion Estimation**. IEEE Transactions on Circuits and Systems for Video Technology, vol. 12, no. 5, May 2002.