

## HEURÍSTICA PARA REDUÇÃO DE TEMPO DE EXECUÇÃO DA ESTIMAÇÃO DE MOVIMENTO DOS MAPAS DE PROFUNDIDADE DO 3D-HEVC

MÁRIO SALDANHA; GUILHERME CORRÊA; MARCELO PORTO;  
LUCIANO AGOSTINI

Universidade Federal de Pelotas – Video Technology Research Group (ViTech)  
*{mrdfsaldanha, gcorrea, porto, agostini}@inf.ufpel.edu.br*

### 1. INTRODUÇÃO

O 3D-High Efficiency Video Coding (3D-HEVC) (SULLIVAN, 2013) é uma extensão do padrão para codificação de vídeos 2D, conhecido como *High Efficiency Video Coding* (HEVC) (SULLIVAN, 2012), para codificação de vídeos 3D. O 3D-HEVC é capaz de fornecer uma redução na taxa de bits para representação de vídeos 3D de aproximadamente 46% e 19% quando comparado com sistemas como *simulcast* HEVC e o *Multiview* HEVC (extensões alternativas do HEVC para codificação de vídeos 3D), respectivamente (TECH, 2016). Um dos fatores principais para essa alta eficiência do 3D-HEVC é a utilização do formato de dados *Multiview Video Plus Depth* (MVD) (MULLER, 2013). O formato MVD utiliza mapas de profundidade juntamente com os quadros de textura durante a codificação, permitindo a síntese de vistas intermediárias (vistas sintetizadas).

Os mapas de profundidade apresentam informações geométricas da cena e fornecem as distâncias da câmera até os objetos. Eles são representados com uma imagem em tons de cinza que é caracterizada por possuir grande regiões homogêneas (interior de objetos e fundo da cena) e arestas bem definidas encontradas nas bordas dos objetos.

O 3D-HEVC adiciona novos algoritmos e estende os algoritmos utilizados no HEVC, tanto para textura quanto para mapas de profundidade. Ainda assim, muitos dos algoritmos usados não exploram as características dos mapas de profundidade, por terem sido originalmente propostos para textura. Então, além de eventuais decisões ruins de codificação, estes algoritmos podem ser ineficientes também em termos de tempo de execução e de consumo de energia.

A predição inter-quadros dos mapas de profundidade no 3D-HEVC é composta pela ferramenta de estimativa de movimento (EM) e a compensação de movimento (CM), herdado do codificador HEVC para vídeos 2D. A EM é uma ferramenta interessante para aplicação de técnicas para redução de tempo de execução, pois demanda um alto esforço computacional e a EM dos mapas de profundidade utiliza um algoritmo complexo desenvolvido para os dados de textura, conhecido como *Test Zone Search* (Tzs) (TANG, 2010).

Neste trabalho é proposta uma heurística capaz de reduzir o tempo de execução e o consumo de energia da EM. A heurística inicialmente classifica o bloco de mapa de profundidade como aresta ou região homogênea, utilizando o algoritmo *Simplified Edge Detector* (SED) (SANCHEZ, 2014). Uma região homogênea pode ser codificada por algoritmos mais rápidos e simples, reduzindo o tempo de processamento e o consumo de energia. Quando o algoritmo SED classifica a região como uma região homogênea, o algoritmo rápido *Iterative-Small Diamond Search Pattern* é aplicado em substituição ao algoritmo Tzs, simplificando a EM. Caso o bloco seja classificado como aresta, nenhuma simplificação é realizada e o codificador segue o fluxo normal da EM utilizando o algoritmo Tzs.

## 2. METODOLOGIA

As Fig. 1 (a) e (b) apresentam duas partes de um quadro de mapa de profundidade para a sequência de vídeo 3D *Shark*. Nestas figuras, o quadrado azul representa o bloco que está sendo codificado e o quadrado em vermelho representa a área de busca utilizada durante a EM. As Fig. 1 (c) e (d) demonstram o mapa de calor destas áreas de busca, considerando o critério de similaridade *Sum of Absolute Differences* (SAD). Regiões em azul escuro denotam os melhores valores de SAD (maior similaridade), enquanto regiões em vermelho representam os piores valores de SAD (menor similaridade).

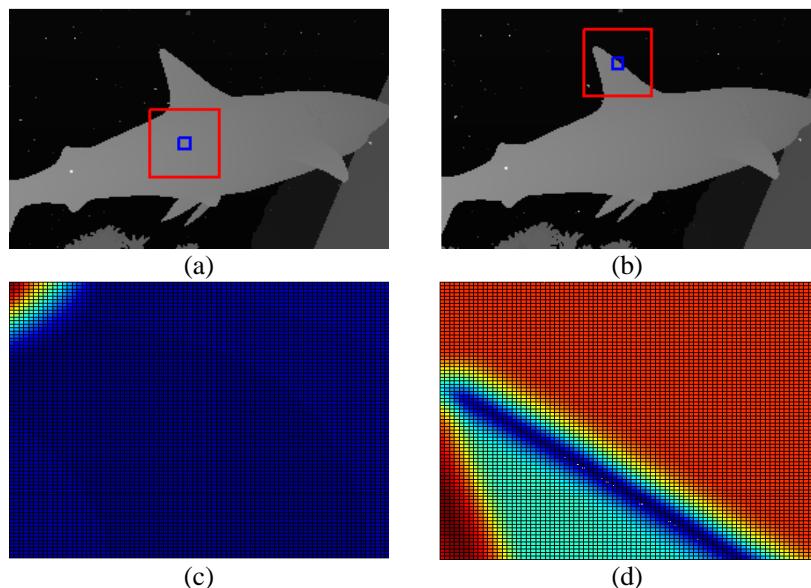


Figura 1 – Exemplo de bloco e a área de busca com regiões (a) homogênea e (b) aresta e o mapa de calor com regiões (c) homogênea e (d) aresta.

Analizando a Fig. 1 é possível notar que quando um bloco que está sendo codificado é composto por região homogênea, o mapa de calor (Fig. 1 (c)) apresenta um comportamento mais simples, com troca de valores mais suaves e com grandes regiões contendo bons valores de SAD próximos ao centro da área de busca. Assim, algoritmos mais leves e rápidos de EM, como o I-SDSP, podem alcançar bons resultados quando aplicados a regiões homogêneas dos mapas de profundidade, reduzindo o número de cálculos de SAD. Por outro lado, quando é analisado o mapa de calor para a região composta por aresta (Fig. 1 (d)), é possível observar um comportamento mais complexo, onde os melhores valores de SAD se encontram em uma região reduzida. Neste caso, um algoritmo mais sofisticado como o TZS é o mais indicado para ser utilizado.

Algoritmos como o I-SDSP necessitam de um número menor de cálculos de SAD quando comparados ao TZS. Reduzir o número de cálculos de SAD é importante, pois além de reduzir o número de operações aritméticas necessárias, também é reduzido o número de acessos à memória, uma vez que uma nova área de busca deve ser carregada da memória a cada novo cálculo de SAD. Tanto os cálculos aritméticos quanto o acesso à memória são muito importantes para definir o tempo de execução e o consumo de energia da EM. Cabe destacar que o acesso à memória é um dos principais elementos que contribui para o consumo de energia nos sistemas digitais atuais. Além disso, a eficiência energética é critério fundamental para codificadores de vídeo rodando em sistemas embarcados como smartphones, câmeras, etc.

### 3. RESULTADOS E DISCUSSÃO

A heurística proposta foi implementada no software de referência do 3D-HEVC, conhecido como 3D-HEVC *Test Model* (3D-HTM), na versão 16.0. As avaliações foram realizadas utilizando as Condições Comuns de Testes (CCT) (RUSANOVSKYY, 2013) do 3D-HEVC. Os resultados obtidos nesta avaliação são comparados com os do 3D-HTM 16.0 original (sem modificações).

A Tabela 1 mostra a eficiência de codificação das vistas sintetizadas utilizando o critério Bjontegaard Delta-rate (BD-rate) e a redução do tempo de codificação obtido pela heurística proposta neste trabalho, considerando o tempo total (textura e profundidade) e somente dos mapas de profundidade. A heurística foi capaz de reduzir, em média, 6,9% do tempo de codificação dos mapas de profundidade e 3,2% do tempo total, com uma perda de 0,148% (em média) na eficiência de codificação (BD-rate).

Tabela 1 – Avaliação da heurística proposta com as CCTs.

Vídeo	BD-rate	Redução tempo de codificação	
		Vistas sintetizadas	Total
Balloons	0,021%	3,6%	7,7%
Kendo	0,025%	3,4%	7,6%
Newspaper_CC	0,106%	3,7%	7,2%
GT_Fly	0,095%	3,3%	7,5%
Poznan_Hall2	0,290%	3,2%	6,8%
Poznan_Street	0,071%	3,6%	7,3%
Undo_Dancer	0,204%	2,3%	5,4%
Shark	0,374%	2,9%	6,0%
<b>Média</b>	<b>0,148%</b>	<b>3,2%</b>	<b>6,9%</b>

A Fig. 2 apresenta a porcentagem de redução de computações de SAD, em comparação com o 3D-HTM utilizando o fluxo tradicional. É possível notar que a heurística reduz mais que 60% dos cálculos de SAD para todas sequências avaliadas e, em média, ocorre uma redução de 68,2% nos cálculos de SAD. Apenas como exemplo, para processar dez segundos de um vídeo de resolução 1920×1088 usando o TZS, são necessários, em média, 1,5 bilhões de cálculos de SAD. Esse resultado é importante, pois os cálculos de SAD demandam um consumo elevado de energia, tanto pelos cálculos em si, que precisam ser realizados em grande número e com frequência elevada, quanto pelo elevado número de acessos à memória que os cálculos de SAD demandam. Estimativas iniciais apontam para uma redução de, pelo menos, 50% no consumo de energia da EM dos mapas de profundidade com o uso da heurística proposta.

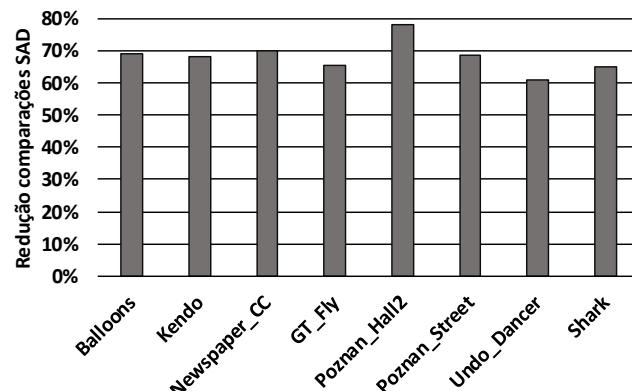


Figura 2 – Redução das comparações de SAD utilizando a solução proposta.

## 4. CONCLUSÕES

Este trabalho apresentou uma heurística para redução de tempo de execução da estimação de movimento dos mapas de profundidade do 3D-HEVC. A heurística proposta teve como objetivo principal reduzir o tempo de execução e o consumo de energia da codificação dos mapas de profundidade, sem inserir perdas significativas na eficiência de codificação. A ideia principal é classificar os blocos do mapa de profundidade em arestas ou regiões homogêneas e aplicar um algoritmo mais simples de estimativa de movimento nas regiões homogêneas. Resultados experimentais foram obtidos utilizando as CCTs para vídeos 3D, e demonstraram uma importante redução no tempo de execução dos mapas de profundidade (6,9% em média), com impacto mínimo na eficiência de codificação das vistas sintetizadas (0,148% em média). Além disso, é obtida uma significativa redução no número de cálculos de SAD (68% em média), contribuindo, também, para a redução do consumo de energia do codificador. Estimativas iniciais apontam para ganhos superiores a 50% no consumo de energia da estimação de movimento dos mapas de profundidade com o uso da heurística proposta. Como trabalhos futuros está planejada a geração de resultados precisos de consumo de energia com o uso da heurística.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

- Müller, K. et al. 3D High-Efficiency Video Coding for multi-view video and depth data. **IEEE Transactions on Image Processing**, vol. 22, n. 9, pp. 3366-3378, Set. 2013.
- RUSANOVSKYY, D.; MULLER, K.; VETRO, A. Common Test Conditions on 3D Core experiments. **Documento: JCT3V-C1100**. Geneva, 2013.
- Sanchez, G. et al. Complexity Reduction for 3D-HEVC Depth Maps Intra-frame Prediction using Simplified Edge Detector Algorithm. **International Conference on Image Processing (ICIP)**, pp. 3209-3213, 2014.
- Sullivan, G. et al. Overview of the high efficiency video coding (HEVC) standard. **IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)**, vol. 22, no. 12, pp. 1649-1668, Dez. 2012.
- Sullivan, G. et al. Standardized Extensions of High Efficiency Video Coding (HEVC). **IEEE Journal of Selected Topics in Signal Processing**, vol. 7, n. 6, pp. 1001-1016, Dez. 2013.
- Tang, X. et al. An Analysis of TZSearch Algorithm in JMVC. **International Conference on Green Circuits and Systems (ICGCS)**, pp. 516-520, 2010.
- 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 (TCSVT)**, v. 26, n. 1, pp. 35-49, Jan. 2016.