

## ESQUEMA PARA REDUÇÃO DA COMPLEXIDADE NA PREDIÇÃO DOS MAPAS DE PROFUNDIDADE DO 3D-HEVC

MÁRIO SALDANHA; BRUNO ZATT; MARCELO PORTO; LUCIANO AGOSTINI

*Universidade Federal de Pelotas  
Grupo de Arquiteturas e Circuitos Integrados  
[{mrdfsaldanha, porto, zatt, agostini}@inf.ufpel.edu.br](mailto:{mrdfsaldanha, porto, zatt, agostini}@inf.ufpel.edu.br)*

### 1. INTRODUÇÃO

Vídeos 3D permitem uma experiência visual com percepção de profundidade da cena, permitindo que os espectadores usufruam de uma visualização diferente da fornecida pelos vídeos 2D (YASAKETHU, 2009). Vídeos 3D são construídos de múltiplas vistas (das múltiplas câmeras), criando a necessidade de codificar um conjunto de vistas simultaneamente. Este processo resulta em um aumento considerável na complexidade de codificação e na largura de banda necessária para transmissão de vídeos 3D, quando comparado aos vídeos 2D. Baseado neste fato, existe uma grande motivação para a investigação e o desenvolvimento de técnicas e algoritmos para a codificação eficiente de vídeos 3D.

O padrão estado da arte para codificação de vídeos 3D é o 3D-*High Efficiency Video Coding* (3D-HEVC) (SULLIVAN, 2013), que é uma extensão do *High Efficiency Video Coding* (HEVC) (SULLIVAN, 2012), que é o estado da arte na codificação de vídeos 2D. O 3D-HEVC adota o formato de dados chamado de *Multi-View plus Depth* (MVD) (MULLER, 2013), que associa um quadro de textura (imagem visualizada pelo usuário final) a um mapa de profundidade correspondente. Mapas de profundidade representam a distância entre os objetos e a câmera. Estes mapas de profundidade possuem uma representação em tons de cinza, composta por valores entre 0 e 255.

Com a utilização do formato MVD ao invés de codificar um grande conjunto de câmeras em um cenário, apenas um subconjunto das câmeras deve ser codificado/transmitido, juntamente com seus mapas de profundidade (capturados por sensores infravermelhos), reduzindo a quantidade de informações para manipular. A partir das informações de textura e dos mapas de profundidade transmitidos, um decodificador 3D-HEVC é capaz de gerar as vistas intermediárias, utilizando técnicas de síntese de vistas, sem a necessidade de transmitir/codificar as vistas intermediárias. É importante destacar, que para uma síntese de vistas eficiente é necessário que as arestas dos mapas de profundidades sejam representadas/codificadas de forma eficiente para não gerar artefatos nas vistas intermediárias.

Na codificação realizada pelo 3D-HEVC, todas as ferramentas utilizadas no HEVC podem ser utilizadas e também foram adicionadas novas ferramentas para explorar as redundâncias presentes em vídeos 3D, como a redundância entre vistas. Além disso, com a necessidade da transmissão dos mapas de profundidade é necessário realizar a codificação destes dados, portanto, ferramentas utilizadas nos dados de textura e novas ferramentas são utilizadas para codificação dos mapas de profundidade.

A predição intra-quadros explora a redundância de informações presente em um mesmo quadro, nesta etapa da codificação dos mapas de profundidades são utilizados modos de predição do HEVC (também utilizados para textura no 3D-HEVC), eficientes em regiões homogêneas, e modos de codificação

desenvolvidos para realizar a codificação eficiente das arestas dos mapas de profundidade, conhecidos como *Depth Modeling Modes* (DMM).

A predição inter-quadros explora a redundância de informações entre quadros temporalmente vizinhos, aplicando um algoritmo de busca para a estimativa de movimento e utilizando o critério de similaridade *Sum of Absolute Differences* (SAD), nesta etapa de codificação dos mapas de profundidade é aplicado o mesmo algoritmo de busca para a estimativa de movimento utilizado para os dados de textura. Entretanto, este processo foi desenvolvido considerando as características dos dados de textura (informações complexas), possuindo uma alta complexidade na codificação dos dados de profundidade que possuem informações mais simples.

Portanto este trabalho tem como objetivo apresentar um esquema composto por dois algoritmos para codificação dos mapas de profundidade: *Simplified Edge Detector* (SED) e *Small Diamond Search Pattern* (SDSP). O primeiro localizado na predição intra-quadros, responsável por identificar quando é necessária a avaliação dos modos DMMs. O SDSP, localizado na predição inter-quadros, é responsável por realizar um processo menos complexo na estimativa de movimento.

## 2. METODOLOGIA

A Fig. 1 (a) apresenta um mapa de variância para blocos de profundidade, onde regiões com azul escuro denotam valores baixos de variância, enquanto regiões em vermelho denotam altos valores de variância. A Fig. 1 (b) apresenta uma função de densidade de probabilidade da variância para a predição intra do HEVC (curva pontilhada vermelha) e os DMMs (curva sólida azul). Analisando a Fig. 1 (b) é possível notar que a predição intra do HEVC é aplicada sobre regiões homogêneas (regiões com baixos valores de variância), enquanto os DMMs são utilizados na codificação de arestas, que consequentemente apresentam altos valores de variância. Portanto, a avaliação dos modos DMMs pode ser descartada quando blocos de profundidade com baixos valores de variância estão sendo codificados.

Com base nisto, o SED é responsável por classificar quando um bloco deve avaliar os modos DMMs ou não, entretanto, esta classificação é realizada com informações apenas das quatro bordas do canto do bloco, sendo menos complexa que o cálculo da variância para o bloco.

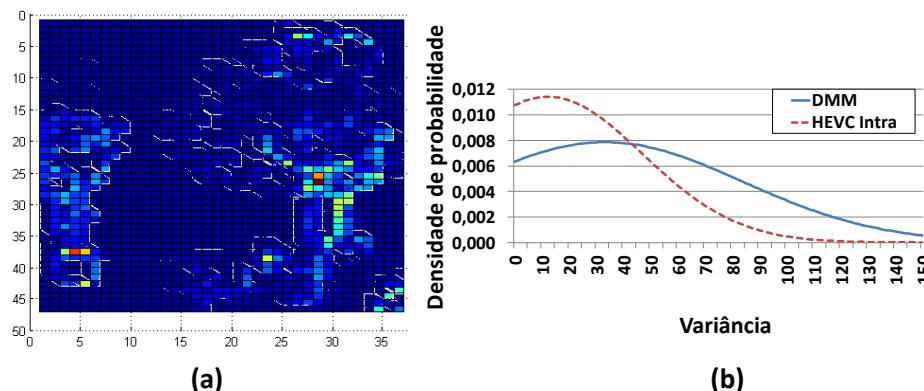


Figura 1 – (a) Mapa de variância e (b) Densidade de probabilidade da variância.

Fig. 2 (a) apresenta uma região da textura, apenas com dados de luminância e (b) seu mapa de profundidade associado. A Fig. 2 (c) e (d) apresentam os

mapas de calor para o SAD para os dados de textura e profundidade correspondentes as regiões acima (Fig. 2 (a) e (b)), respectivamente. Regiões em azul escuro denotam um bom resultado para a estimativa de movimento, enquanto regiões em vermelho representam um resultado ruim para a estimativa de movimento. É possível notar na Fig. 2 (c) que o mapa de calor para textura possui um comportamento complexo, com variações abruptas de valores. Enquanto, os mapas de profundidade possuem um comportamento mais simples, com grandes regiões homogêneas e variações de valores mais suaves.

Portanto, um algoritmo sofisticado necessita ser utilizado para os dados de textura, devido a seu comportamento complexo. No software de referência do 3D-HEVC é utilizado o algoritmo TZ Search (TZS). No entanto, um algoritmo mais simples, como o SDSP, pode ser utilizado para a codificação dos mapas de profundidade, alcançando uma boa eficiência na codificação e reduzindo a complexidade computacional, quando comparado ao TZS.

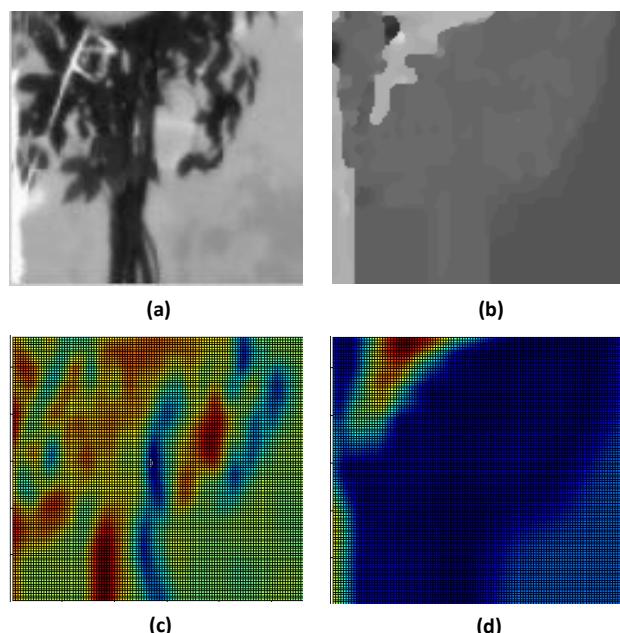


Figura 2 – (a) Textura e (b) profundidade. Mapas de calor do SAD (c) textura (d) profundidade.

### 3. RESULTADOS E DISCUSSÃO

A Tabela 1 apresenta os resultados do esquema proposto aplicado no HTM-10.2. Nesta avaliação foram utilizadas as Condições Comuns de Testes (CCT) (RUSANOVSKYY, 2013). Na Tabela 1 a coluna “Redução SAD”, representa a redução de cálculos de SAD (blocos comparados) durante o processo da estimativa de movimento na predição inter-quadros aplicando a solução proposta. Os resultados apresentados na coluna “Vistas Sintetizadas (BD-rate)” representam o acréscimo na taxa de bits necessária para representação das vistas sintetizadas (vistas geradas através da síntese dos dados de textura e profundidade) para uma mesma qualidade do vídeo. A coluna “Redução Tempo Total” apresenta a redução do tempo de codificação de todo codificador. A coluna “Redução Tempo Profundidade” representa a redução do tempo de codificação considerando apenas os dados de profundidade.

Analizando os resultados, o esquema proposto foi capaz de reduzir em média 88,69% os cálculos de SAD, quando comparado ao TZS. Esta redução na comparação dos blocos ocasiona em uma redução do tempo de codificação de

em média 22,3%, considerando apenas os mapas de profundidade, que é o foco da solução. Considerando o tempo de codificação total (textura e profundidade) foi possível alcançar uma redução de 12,8%, em média. Este ganho expressivo na redução da complexidade de codificação dos mapas de profundidade causa um impacto de apenas 0,69%, em média, nas vistas sintetizadas.

Tabela 1 – Resultados da avaliação do esquema proposto no HTM-10.2.

Vídeos	Redução SAD	Vistas Sintetizadas (BD-rate)	Redução Tempo Total	Redução Tempo Profundidade
Balloons	89,37%	0,22%	13,9%	23,6%
Kendo	91,34%	0,33%	14,4%	24,5%
Newspaper_CC	91,27%	1,47%	15,8%	25,1%
GT_Fly	87,72%	0,59%	11,3%	22,0%
Poznan_Hall2	87,38%	1,00%	12,7%	20,6%
Poznan_Street	88,46%	0,33%	11,4%	21,0%
Undo_Dancer	85,29%	0,92%	9,9%	19,1%
Média	<b>88,69%</b>	<b>0,69%</b>	<b>12,8%</b>	<b>22,3%</b>

#### 4. CONCLUSÕES

Neste trabalho foi apresentado um esquema para redução da complexidade de codificação dos mapas de profundidade, focando nas previsões intra e inter-quadros. Onde foi utilizado o algoritmo SED para classificar se um bloco necessita realizar a avaliação dos modos DMMs ou pode evitar esta avaliação realizando uma simplificação na etapa da previsão intra-quadros. Também foi utilizado o algoritmo SDSP para reduzir a complexidade do algoritmo de busca utilizado na estimação de movimentos da previsão inter-quadros. A solução é capaz de reduzir consideravelmente o tempo de codificação com um impacto mínimo na eficiência da codificação. Além disso, o SDSP é capaz de reduzir significativamente o número de blocos comparados, quando comparado ao TZS. Esta redução é uma característica desejável para desenvolvimento em *hardware*, pois reduz o número de acessos a memória.

#### 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.

Sullivan, G. et al.. Overview of the high efficiency video coding (HEVC) standard. **IEEE Trans. Circuits Syst. Video Technol.**, 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.

Yasakethu, S.L.P. et al.. Analyzing Perceptual Attributes of 3D Video. **IEEE Transactions on Consumer Electronics**, pp. 864 - 872, Ago. 2009.