

ANÁLISE ESTATÍSTICA SOBRE O USO DA PREDIÇÃO INTRA QUADRO EM MAPAS DE PROFUNDIDADE NO PADRÃO EMERGENTE 3DVC

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

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

1. INTRODUÇÃO

As aplicações 3D estão ganhando espaço nos últimos anos. Isto está ocorrendo porque o custo de equipamentos que gravam e reproduzem vídeos 3D está diminuindo e está aumentando o interesse em pesquisa e desenvolvimento de novas tecnologias para codificar vídeos 3D de forma mais eficiente.

Mapas de profundidade trazem informação adicional aos vídeos tradicionais (textura) e sua codificação está inserida ao padrão emergente 3DVC (3D-*High Efficiency Video Coding*) (TECH, 2013) (MULLER, 2013). Mapas de profundidade são utilizados na codificação de vídeos 3D onde cada canal de textura está associado a um canal de profundidade. Estes canais de profundidade são representados por imagens monocromáticas que indicam a distância entre o objeto capturado e a câmera.

A vantagem no uso de canais de profundidade associados a canais de textura é que é possível gerar os vídeos de canais intermediários (de textura) de forma eficiente através da interpolação de alguns canais de textura e profundidade (MERKLE, 2007). Isto reduz drasticamente a necessidade de processamento para codificação de vídeo e armazenamento e/ou largura de banda para transmissão do vídeo. Por exemplo: considerando um vídeo com nove câmeras e seus canais de profundidade associados, podem ser codificadas as câmeras um, cinco e nove mais os seus canais de profundidade. Com isso os outros canais de textura desejados podem ser sintetizados e é possível obter a redução de 55,6% em *bitrate* (ao invés de codificar nove canais de textura se codifica três canais de textura e três de profundidade).

A Figura 1 (a) apresenta um canal de textura e a Figura 1 (b) seu mapa de profundidade associado. Nos mapas de profundidade é possível identificar claramente os objetos presentes na imagem. Isto é uma característica importante dos mapas de profundidade, eles possuem várias regiões homogêneas e arestas bem definidas.



Figura 1 – (a) Imagem de textura (b) mapa de profundidade associado.

A predição intra-quadro tradicional para vídeos de textura pode gerar ruído ao fazer a codificação de mapas de profundidade. Baseado nisso, o padrão emergente 3DVC utiliza novos modos de predição intra para mapas de profundidade. No caso, para fazer a codificação são testados os modos intra tradicionais e, além disso, os novos modos de codificação intra de profundidade. Técnicas para simplificação da codificação de mapas de profundidade devem ser estudadas a fim de conseguir melhorar o compromisso entre a eficiência da codificação e a complexidade computacional.

Este artigo apresenta um estudo sobre a complexidade computacional da codificação de mapas de profundidade e uma análise estatística sobre a escolha do modo de codificação intra de um bloco de acordo com características do bloco que está sendo codificado.

2. METODOLOGIA

Foram codificados sete vídeos 3D com mapas de profundidade associados e foi medido o tempo utilizado durante a codificação de quadros de textura e quadros de profundidade. Os resultados são apresentados na Tabela 1 onde é mostrado o tempo utilizado na codificação de textura, de profundidade e a porcentagem do tempo gasto na codificação de profundidade em relação ao tempo total do codificador.

Através da Tabela 1 pode-se notar que a complexidade computacional de profundidade representa em média 25,28% da complexidade computacional de todo o codificador, ou seja, adicionar a codificação dos canais de profundidade acrescentou uma complexidade de 33,83% ao tempo total de um codificador que previamente apenas codificava textura.

Tabela 1 – Divisão de tempo no codificador entre profundidade e textura.

Vídeo	Tempo textura	Tempo Profundidade	% de tempo usado em profundidade
Balloons (1024x768)	30.524	11.599	27,54%
Kendo (1024x768)	34.580	13.692	28,36%
Newspaper (1024x768)	29.558	11.514	28,03%
GT_Fly (1920x1088)	69.693	23.216	24,99%
Poznan_Hall2 (1920x1088)	47.505	13.701	22,39%
Poznan_Street (1920x1088)	55.098	19.862	26,50%
Undo_Dancer (1920x1088)	64.033	18.404	22,32%
Média	47.284	15.998	25,28%

A codificação de vídeos de alta resolução em geral necessita de uma alta complexidade computacional e muitas vezes os dispositivos tem grande dificuldade de conseguir fazer a codificação em tempo real sem necessitar de uma grande energia consumida. Este aumento da complexidade computacional se torna algo problemático visto que irá ocasionar um atraso na codificação e, além disso, será necessário um novo circuito integrado para fazer a codificação deste canal de profundidade.

Através desta análise aliada ao fato de que a codificação de profundidade é um artifício novo da codificação de vídeo do padrão emergente 3DVC é um desafio motivador o fato de que a complexidade computacional da codificação de mapas de profundidade ocupa um grande espaço da codificação de vídeos 3D e

possui, até o momento, poucos algoritmos com a capacidade de reduzir esta complexidade.

Uma característica que difere muito quadros de profundidade de quadros de textura é a presença de regiões muito homogêneas e para separar estas regiões arestas com o gradiente bastante abrupto. Com isso, foi proposto um algoritmo para identificar se o bloco representa uma região de aresta. Este algoritmo compara os quatro pixels dos extremos de um bloco e caso a maior diferença entre eles seja maior do que um valor pré-definido, então é dito que este bloco representa uma aresta. Caso contrário, é dito que o bloco representa uma área homogênea.

Foram codificados os sete vídeos novamente e para cada bloco de profundidade codificado foi armazenado o modo de predição selecionado, intra tradicional ou utilizando os novos modos. Estes dados foram utilizados para gerar a densidade de probabilidade da escolha de modos intra tradicionais ou novos modos com relação à maior diferença entre os extremos dos blocos. Esta análise permite que seja definido o valor utilizado no algoritmo previamente mencionado.

3. RESULTADOS E DISCUSSÃO

Os experimentos foram divididos em duas classes. A primeira analisou apenas os resultados para vídeos de resolução HD 1080p e a segunda para vídeos com resolução 1024x768 pixels. Até o momento, a análise foi feita apenas para blocos de tamanho 32x32 (modos intra podem ser codificados em blocos 4x4, 8x8, 16x16 e 32x32).

A Figura 2 apresenta a densidade de probabilidade da escolha entre os novos modos e intra tradicional para vídeos de resolução HD 1080p em função da maior diferença entre os pixels das extremidades do bloco. A Figura 3 apresenta o mesmo para vídeos de resolução 1024x768 pixels.

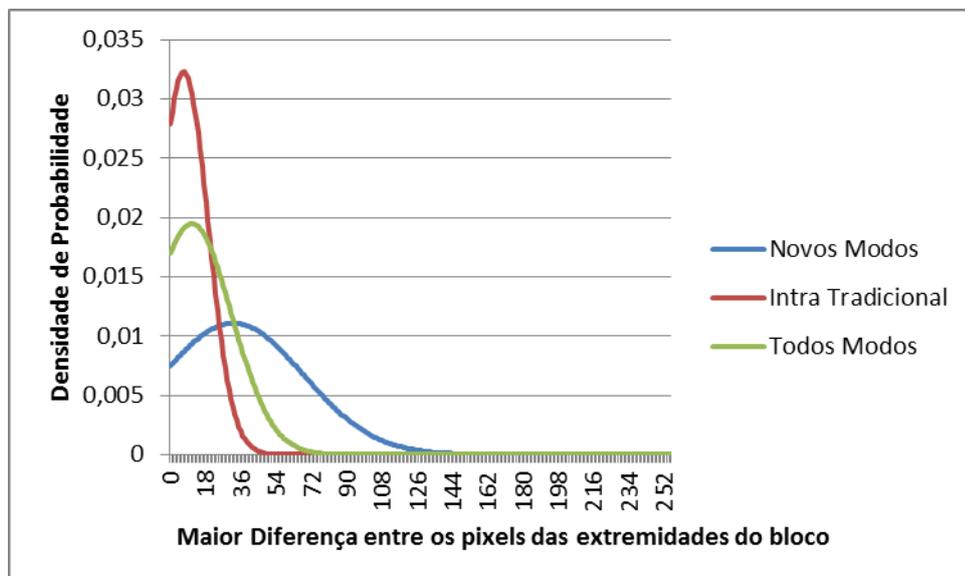


Figura 2 – Estatística para vídeos HD 1080p – tamanho de bloco 32x32

Em ambas figuras, é possível notar que após um determinado valor, a probabilidade da escolha de modos intra tradicionais é praticamente nula, enquanto ainda existe uma grande quantidade de blocos que são avaliados. Tradicionalmente no padrão emergente 3DVC, são avaliados os novos modos e os modos intra tradicionais sem utilizar um critério para fazer a simplificação.

Com esta análise estatística pode-se notar claramente que ao identificar se um bloco é uma aresta, a codificação utilizando a predição intra tradicional é totalmente indesejada e não é necessário gastar tempo de processamento com isto. Neste caso, pode-se cortar a predição intra tradicional nos valores em que a densidade probabilidade de da escolha de intra tradicional seja baixa.

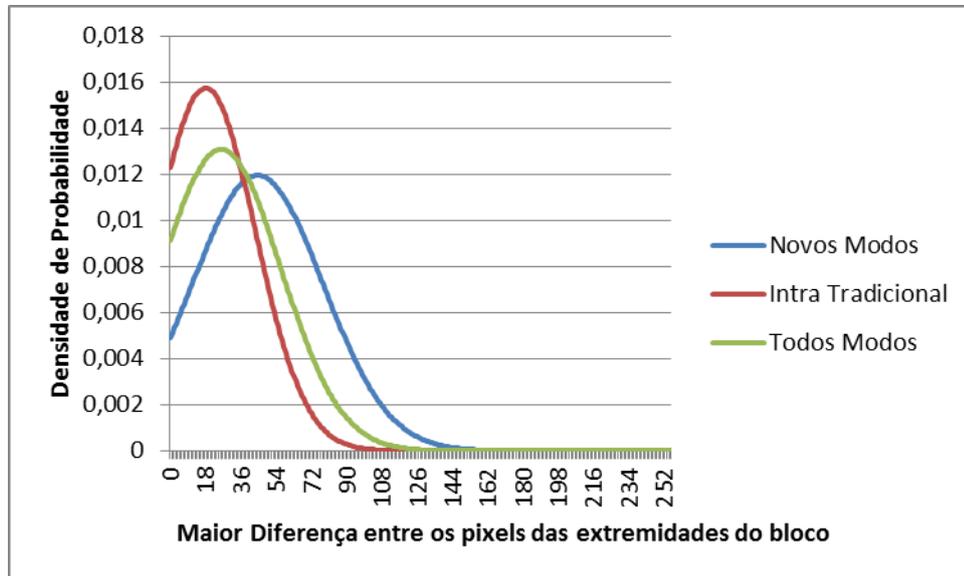


Figura 3 – Estatística baixa resolução – tamanho de bloco 32x32

4. CONCLUSÕES

Este trabalho apresentou um estudo sobre a complexidade da adição de quadros de profundidade na codificação de vídeos 3D. Após, foi apresentada uma análise estatística com o intuito de permitir decidir se um bloco de profundidade representa uma região homogênea ou uma aresta.

Com a análise estatística foi possível concluir que ao analisar a maior diferença entre os quatro pixels dos extremos de um bloco é possível determinar se o bloco deve ou não ser codificado com a predição intra tradicional.

Como trabalhos futuros, pretende-se aplicar este conhecimento na elaboração de um algoritmo focado no padrão emergente 3DVC com a finalidade de decidir se um bloco deve ou não ser codificado com os modos intra tradicionais. Com isto é possível realizar uma grande simplificação na codificação de mapas de profundidade sem inserir uma perda significativa na qualidade da codificação. Além disso, pretende-se fazer uma análise estatística completa para os outros tamanhos de blocos disponíveis para deixar esta análise mais robusta.

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.