

AVALIAÇÃO DA DECISÃO DE MODO DO CODIFICADOR DE VÍDEO 3D-HEVC

RUHAN CONCEIÇÃO; **GUILHERME CORRÊA**; **MARCELO PORTO**;
BRUNO ZATT; **LUCIANO AGOSTINI**

Universidade Federal de Pelotas - Grupo de Arquiteturas e Circuitos Integrados
 {radconceicao,gcorrea,porto,zatt,agostini}@inf.ufpel.edu.br

1. INTRODUÇÃO

Nos últimos anos vem ocorrendo uma crescente popularização de vídeos digitais e dos dispositivos capazes de manipularem os mesmos, tais como *smartphones*, *tablets*, aparelhos televisores, *Blu-ray players*, *video games*, etc. Apesar da popularização dos vídeos digitais, existem diversos desafios associados a estes, destacando-se reduzir o volume de dados necessários para representar os vídeos através de técnicas eficientes de codificação.

Este desafio torna-se maior ao considerarmos vídeos de três dimensões (3D), os quais são gravados utilizando mais de uma câmera – ou seja, mais de uma vista – além de possuir um mapa de profundidade associado a cada uma destas vistas. Os mapas de profundidade são imagens monocromáticas que armazenam a distância da câmera até os elementos da imagem.

A Figura 1 apresenta uma imagem extraída de um vídeo 3D sem codificação alguma denominado *Balloons* (TANIMOTO, 2010). A imagem é composta pela textura (a) acompanhada pelo seu respectivo mapa de profundidade (b).



Figura 1. Vista de uma câmera 3D, (a) textura e (b) mapa de profundidade associado

Para demonstrar a necessidade das técnicas de compressão, vamos considerar um vídeo 3D de alta resolução HD (*High Definition*) 1080p (1920x1080 pixels), o qual foi gravado utilizando-se três câmeras. Para armazenar duas horas deste vídeo, seria necessário 391 *Gigabytes* em disco. Caso o mesmo fosse transmitido por um canal de comunicação, seria necessário que este fosse capaz de transmitir 55,62 *Megabytes* por segundo. Assim, destaca-se a necessidade de comprimir este vídeo através de técnicas de codificação.

Atualmente, um novo padrão de codificação de vídeos em três dimensões está sendo desenvolvido, o *3D High Efficiency Video Coding* (3D-HEVC) (MULLER, 2013). Este padrão emergente é baseado no padrão recém lançado para codificação de vídeos denominado *High Efficiency Video Coding* (HEVC) (JCT-VC, 2013).

Ao codificar um vídeo, o 3D-HEVC (assim como o HEVC) divide cada imagem em blocos de tamanho 64x64 pixels denominados *Coding Tree Units* (CTUs). Cada CTU é composta por uma ou mais *Coding Units* (CUs), as quais podem ser de tamanho 64x64, 32x32, 16x16 ou 8x8. Basicamente, cada CU

define para o conjunto de pixels qual tipo de predição utilizar (*intra- ou inter-quadro*), o tipo de particionamento em *Prediction Unit* (PU), etc (SULLIVAN, 2012).

A Figura 2 apresenta um exemplo de particionamento das CUs (linhas pretas) e PUs (linhas brancas) de um quadro do vídeo *Balloons*. As CUs vermelhas indicam que a predição utilizada foi a *intra-quadros*, ao passo que as azuis indicam que a predição selecionada foi a *inter-quadros*. Por fim, as CUs de cor cinza indicam que o modo de predição utilizado foi *skip*. O modo de predição *skip* é um caso especial do modo de predição *merge*, o qual infere as informações de codificação, reduzindo consideravelmente a informação a ser armazenada. O modo de predição *skip* apresenta ganhos de compressão ainda superiores, visto que este além de inferir a forma em que é feita codificação do bloco, infere também as amostras do mesmo, não armazenando informação residual (SULLIVAN, 2012). Além das informações elucidadas acima, a Figura 2 também destaca uma CTU da imagem.

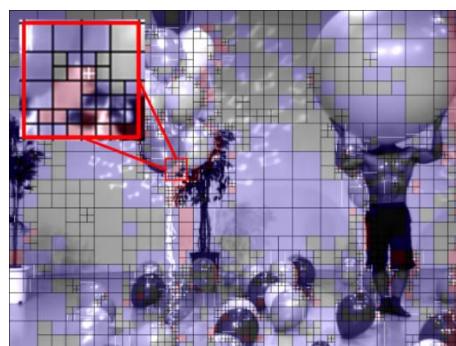


Figura 2. Particionamento de CUs e PUs de uma vista de textura do vídeo *Balloons*

Considerando estas informações, este trabalho visa apresentar um estudo sobre a decisão de modo (*Mode Decision - MD*) do codificador 3D-HEVC. O MD é responsável por definir os tamanhos de CUs, os modos de predição, qual PU e qual particionamento utilizar, etc. Este estudo é um estágio inicial de pesquisa sobre redução de complexidade no codificador 3D-HEVC, a qual visa reduzir o tempo de codificação sem ônus significativos na qualidade de imagem e tamanho do *bitstream* gerado.

2. METODOLOGIA

Visando elaborar o estudo sobre o MD do 3D-HEVC, primeiramente, o software de referência do padrão, o *HEVC Test Model 14.1* (HTM 14.1) (FRAUNHOFER, 2015), foi investigado. Posteriormente, foram investigadas as funções responsáveis pela decisão de modo do 3D-HEVC no *software*, inserindo trechos de código que realizassem a extração dos dados utilizados.

Uma vez realizada a extração dos dados, foi desenvolvido um programa em C# capaz de interpretar as informações geradas pela a extração dos dados. Assim, viabilizando o estudo da MD do 3D-HEVC.

Neste trabalho, foram avaliados oito quadros do vídeo *Balloons*, utilizando-se parâmetro de quantização (*Quantization Parameter – QP*) 25 para quadros de textura e 30 para mapas de profundidade. Todo o estudo considerou, para cada *Coding Unit*, o seu tamanho, se o codificador optou por efetuar *merge* ou *skip*, o tipo de predição (*intra* ou *inter*) e o particionamento da PU (simétrico ou assimétrico). Ao todo, aproximadamente 63.000 CUs foram analisadas neste trabalho.

3. RESULTADOS E DISCUSSÃO

A primeira análise foi realizada visando quantificar o número total de CUs, de cada tamanho, com predição *intra* e *inter-quadros*, considerando quatro diferentes cenários: CU pertence (ou não) a um quadro do tipo I (TI), e se a CU pertence (ou não) a um mapa de profundidade (MP). A Tabela 1 apresenta os dados totais desta análise, onde MP:0 significa quadro de textura e MP:1 mapa de profundidade. Além disso, TI:0 significa que a CU não pertence a um quadro de tipo I, ao passo que TI:1 significa o contrário.

Por fim, a tabela também apresenta a porcentagem de amostras do cenário que pertencem a uma CU com a dada configuração. É importante salientar que CUs maiores compreendem mais amostras; assim, uma CU 64x64 equivale a 64 CUs de tamanho 8x8. Desta forma, a porcentagem leva em consideração o número de amostras abrangidas pela CU.

Tabela 1. MD do 3D-HEVC observando tamanho de CUs e tipo de predição

		MP:0 TI:0	%	MP:0 TI:1	%	MP:1 TI:0	%	MP:1 TI: 1	%
P-Intra	64x64	4	0,04	5	2,6	386	4,28	58	30,21
	32x32	96	0,27	252	32,81	1.013	2,81	204	26,56
	16x16	364	0,25	1.029	33,5	2.447	1,69	745	24,25
	8x8	884	0,15	3.820	31,09	5.111	0,88	2.332	18,98
P-Inter	64x64	6.436	71,32	0	0	7.163	79,38	0	0
	32x32	7.580	21	0	0	3.184	8,82	0	0
	16x16	8.332	5,77	0	0	2.691	1,86	0	0
	8x8	6.892	1,19	0	0	1.585	0,27	0	0

Como visto, nenhuma CU foi definida com predição *inter-quadros* ao tratar-se de um quadro I. Isto é esperado, visto que estes quadros são codificados apenas por CUs com predição *intra*. Além disso, observa-se que nos demais quadros a porcentagem de CUs com predição *intra* é muito inferior às CUs com predição *inter*. Isto deve-se ao fato que a codificação *inter-quadros* apresenta ganhos muito mais relevantes que a codificação *intra-quadros*.

Outra observação a ser feita na Tabela 1, considerando quadros do tipo I, é o aumento no número de CUs de tamanho 64x64 para mapas de profundidade. Isto se deve ao fato de que os mapas de profundidade possuem regiões mais homogêneas, as quais viabilizam o uso de CUs maiores na codificação *intra*.

Além da análise acima, observou-se também o número de CUs que foram codificadas com *merge* e *skip*, além do número de CUs com particionamento quadrático e assimétrico. Esta análise foi realizada considerando de forma isolada os quadros de textura e mapas de profundidade, além dos tamanhos de CUs. Os resultados da análise são apresentados na Tabela 2.

Como é possível perceber, tanto em quadros de textura quanto em mapas de profundidade, a porcentagem de CUs do tipo *skip* e *merge* cai conforme o tamanho da CU diminui. Além disso, observa-se que a porcentagem de CUs *merge* e *skip* são menores em mapas de profundidade quando comparado com as CUs de quadros de textura. Outro dado observado na Tabela 2 é a porcentagem pouco expressiva de PUs assimétricas utilizadas na codificação de mapas de profundidade. O porquê destes fatos será alvo de estudo no decorrer desta pesquisa.

Tabela 2. MD do 3D-HEVC observando taxa de ocorrência de *merge* e *skip*, e particionamento de PUs

Modo	Textura				Mapa de Profundidade			
	64x64	32x32	16x16	8x8	64x64	32x32	16x16	8x8
Total de CUs	6.440	7.676	8.996	7.776	7.549	4.197	5.138	6.696
Merge (%)	97,9	91,7	84,7	78,7	89,8	66,74	45,82	21,2
Skip (%)	93,1	73,65	58,3	49	83,7	54,3	35,5	17,4
PU Quadrada (%)	97,4	92,3	91,3	95,7	99,7	98,5	98,7	99,8
PU Assimétrica (%)	2,6	7,7	8,7	4,3	0,3	1,5	1,3	0,2

4. CONCLUSÕES

Este trabalho apresentou uma avaliação da decisão de modo do codificador de vídeos 3D-HEVC, demonstrando quais são as abordagens mais frequentemente utilizadas pelo codificador em determinados cenários.

As avaliações presentes neste trabalho integram uma linha pesquisa em codificação de vídeos, a qual visa definir estratégias de redução de complexidade do codificador 3D-HEVC. Estas estratégias devem reduzir o tempo de codificação, sem deteriorar a eficiência do codificador em termos de qualidade de compressão e de imagem.

Os principais resultados da avaliação deste trabalho apontaram que mapas de profundidade tendem a utilizar mais as CUs de tamanho 64x64, quando comparado com os quadros de textura. Além disso, embora o já codificador apresente baixa utilização de PUs assimétricas para quadros de textura, esta taxa diminui ainda mais ao considerarmos mapas de profundidade.

Como trabalho futuro, considerando a tarefa de análise de codificação, espera-se observar o comportamento do codificador 3D-HEVC ao codificar outros vídeos que apresentem características de imagem e resoluções diferentes.

5. REFERÊNCIAS BIBLIOGRÁFICAS

TANIMOTO Laboratory. **Balloons**. Departamento of Information Electronics, Nagoya University, 2010. Acessado em 07 jul. 2015. Online. Disponível em: <http://www.tanimoto.nuee.nagoya-u.ac.jp/~fukushima/mpegftv/yuv/Balloons/>

MULLER, K. 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.

SULLIVAN, G. J. Overview of the High Efficiency Video Coding (HEVC) Standard. **IEEE Transactions on Circuits and Systems for Video Technology**. v.22, n.12, p. 1649 – 1668, 2012.

JCT-VC. High Efficiency Video Coding text specification draft 10, doc. JCTVC-L1003, Genebra, Suíça, 2013.

FRAUNHOFER, Heinrich Hertz Institute. **HTM software repository**. Acessado em 07 jul. 2015. Online. Disponível em: https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/