

IMPACTO DAS FERRAMENTAS DE PREDIÇÃO INTER-QUADROS NOS MAPAS DE PROFUNDIDADE DE VÍDEOS 3D NO PADRÃO 3D-HEVC

ÂNDRIO FERREIRA DE ARAÚJO CAMPOS; RUHAN CONCEIÇÃO; BRUNO ZATT, MARCELO PORTO, LUCIANO VOLCAN AGOSTINI

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

1. INTRODUÇÃO

Com o passar dos anos, dispositivos eletrônicos com suporte a vídeos digitais, como por exemplo, *smartphones* e *tablets*, são cada vez mais fáceis de serem encontrados. Além disso, a qualidade e a resolução desses vídeos estão cada vez maiores, demandando a necessidade de representar um volume de dados cada vez maior. Com isso, a busca para aperfeiçoar os codificadores de vídeos se tornou uma atividade de grande importância, pois dispositivos que processam vídeos digitais devem ser capazes de processar vídeos em alta resolução em tempo real.

No cenário atual, as aplicações 3D vêm ganhando espaço e junto com elas, o aumento no interesse em pesquisas e desenvolvimento de novas tecnologias eficientes para codificação de vídeos 3D.

Em vídeos 3D, uma cena é capturada simultaneamente por várias câmeras em posições diferentes, o que gera uma grande quantidade de informações a serem processadas. A codificação de vídeos tem como objetivo explorar características redundantes dentro do vídeo para reduzir o seu volume de informações. Um modo de reduzir informações redundantes é utilizando a predição inter-quadros (Silva, 2013), onde informações de quadros vizinhos são utilizados para codificar o quadro atual.

O padrão emergente de codificação de vídeos 3D, o 3D-HEVC (MULLER, 2013), é baseado no padrão HEVC – *High Efficiency Video Coding* (ITU-T, 2013). Dentro das suas principais características está a introdução do conceito de mapas de profundidade onde cada quadro de textura é associado a um mapa de profundidade (Sanchez, 2014), estes mapas são representados utilizando tons de cinza para representar a distância entre a câmera e os objetos. A Figura 1 contém um quadro de textura (a) com seu respectivo mapa de profundidade (b).



(a)



(b)

Figura 1. Quadro de textura (a) e seu quadro de profundidade (b).

O processo de codificar um vídeo envolve a utilização de diversas ferramentas, o que torna o processo custoso. Além disso, quando este cenário envolve vídeos 3D, este custo é ainda maior devido ao aumento do número de câmeras e a introdução de novas ferramentas. Entre as novas ferramentas, destaca-se a integração dos mapas de profundidades e de suas ferramentas específicas.

Visando reduzir o custo do processo de codificação de vídeos, este trabalho propõe uma discussão sobre o impacto das ferramentas de predição inter-quadros aplicadas nos mapas de profundidade.

2. METODOLOGIA

Para o desenvolvimento deste trabalho, foi necessário realizar um estudo sobre os mapas de profundidade, após utilizando o software de referência *HEVC Teste Model 14.1* (HTM 14.1) (3D-HEVC-HTM, 2015) foi feita uma análise sobre a predição inter-quadros e sobre os modos inter-quadros presentes no codificador 3D-HEVC. Finalmente, foi feito uma análise baseada em testes de simulações sobre o impacto da predição inter-quadros aplicada nos mapas de profundidade em diferentes cenários de codificação.

Primeiramente é necessário ter conhecimento de que durante a codificação o vídeo a ser codificado é dividido em vários quadros, após, estes quadros são particionados em blocos menores, para que cada um destes blocos passe, individualmente, por um processo completo de codificação, estes blocos menores são chamados de *Coding Units* (CU) (ITU-T, 2013).

As *Coding Units* são muito utilizadas nas predições inter-quadros. Este modulo é responsável por encontrar informações redundantes presentes entres quadros temporalmente vizinhos, e também em quadros de outras câmeras (considerando aplicações 3D). Este modo de predição é muito eficaz em compressão de dados devido à grande semelhança entre os quadros vizinhos. O modulo de predição inter-quadros utilizado no codificador 3D-HEVC, assim como no HEVC, particiona estas *Coding Units* em *Prediction Units*, as quais podem ter diferentes formatos como pode ser observado na Figura 2.

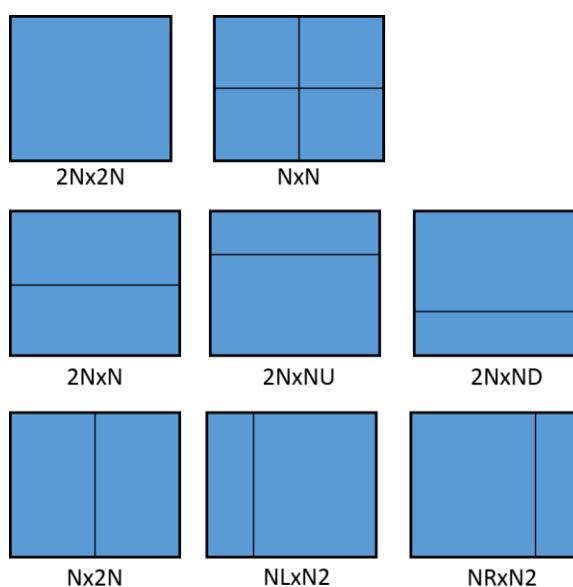


Figura 2. Formatos de *Prediction Units*.

Neste trabalho, três cenários diferentes de codificação foram avaliados. Estes cenários foram selecionados levando em consideração, a frequência em que os modos são utilizados pela predição inter-quadros em mapas de profundidade e por seu ganho em compressão de dados.

A primeira avaliação feita foi analisar o comportamento do codificador quanto a restrição no uso dos modos de predição inter-quadros $N \times N$, $2N \times N$ e $N \times 2N$. Estes modos são muito aplicados em bordas de objetos, porém, a ocorrência de PUs com estes formatos é muito pequena.

O segundo cenário foi desabilitar o modo *Skip*. Este modo infere tanto a forma como o bloco será codificado, como as amostras do bloco a ser codificado. O *Skip* não armazena nenhuma informação residual, o que gera um grande ganho de compressão.

O último cenário analisa a restrição conjunta dos modos *Skip* e $2N \times 2N$, os quais são os modos mais utilizados na predição inter-quadros além de terem o maior ganho em compressão de dados.

3. RESULTADOS E DISCUSSÃO

Para a realização da análise proposta neste trabalho, foi observado o impacto das ferramentas de predição inter-quadros aplicadas nos mapas de profundidade em diferentes cenários de codificação. Os testes foram avaliados em termos de qualidade de imagem (Y-PSNR) e de taxa de bits (Bitrate).

Para avaliações dos testes de cenários, primeiramente o modulo de predição inter-quadros foi avaliado sem nenhuma modificação, ou seja, foram feitas simulações utilizando o codificador de vídeo sem nenhuma alteração, com o objetivo de ser utilizadas como parâmetro de comparação.

A Tabela 1 apresenta os resultados obtidos nos testes de simulações, onde são apresentadas duas sequências de mapas de profundidade para permitir análises mais precisas, cada sequência de mapas de profundidade, são diversos quadros de mapas de uma câmera.

Tabela 1. Resultados de simulação.

Primeira Sequência				
	Original	Teste 1	Teste 2	Teste 3
Y-PSNR (dB)	39,16	-0,05	0,38	-1,21
Bitrate (bits/seg)	133,65	-1,93	38,1	194,9
Segunda Sequência				
	Original	Teste 1	Teste 2	Teste 3
Y-PSNR (dB)	38,21	-0,12	0,4	-0,9
Bitrate (bits/seg)	161,71	0,18	47,1	244

Comparando os três diferentes testes, com o cenário original do codificador, é possível verificar que o primeiro teste (Teste 1) teve um impacto muito pequeno no processo de codificação, isso se dá pelo fato de que os modos que não foram executados ($N \times N$, $2N \times N$ e $N \times 2N$) são muito pouco utilizados pela predição inter-quadros.

Observando os resultados do Teste 2, é possível notar que a qualidade da imagem (Y-PSNR) aumentou, isso é consequência da retirada do modo *Skip*, isto forçou uma busca dentro do bloco sobre o melhor modo inter-quadros possível.

Consequentemente, a taxa de bits também aumentou, isso se dá pela perda do ganho em compressão proporcionada pelo modo *Skip*.

O Teste 3 foi o cenário com o maior impacto, não só a qualidade do vídeo caiu como a taxa de bits teve aumento muito expressivo. Este resultado é a consequência de restringir o codificador a não utilizar os dois modos mais comuns de predição inter-quadros, obrigando a processar o quadro com outros modos menos utilizados.

4. CONCLUSÕES

Este trabalho apresentou uma análise do impacto das ferramentas de predição inter-quadros aplicadas nos mapas de profundidade na codificação de vídeos 3D no padrão 3D-HEVC. A predição inter-quadros é o módulo de predição com o maior ganho em compressão de dados de um codificador de vídeo, isto faz com que pesquisas a fim de otimizações para este módulo sejam essenciais.

Observando os resultados, é possível analisar como uma pequena alteração na predição inter-quadros dentro dos mapas de profundidade, sendo a restrição do uso de ferramentas pouco ou muito utilizadas, gera um grande impacto tanto na qualidade quanto na taxa de bits de um vídeo, isto motiva a pesquisar e otimizar este módulo.

Como trabalho futuro, será dada continuação às pesquisas sobre a predição inter-quadros do codificador de vídeo 3D-HEVC, com o objetivo de gerar otimizações que possam reduzir o custo computacional da predição com pequeno impacto na qualidade.

5. REFERÊNCIAS BIBLIOGRÁFICAS

SILVA, M. **Heurísticas para Redução da Complexidade na Predição Inter-Quadros do Padrão de Codificação de Vídeo Emergente HEVC**. 2013. Trabalho de Conclusão de Curso (Graduação em Ciência da Computação), Universidade Federal de Pelotas.

MULLER, K.; et al. 3D High-Efficiency Video Coding for Multi-View Video and Depth Data. **IEEE Transactions on Image Processing**, p.3366-3378, 2013.

ITU-T, International Telecommunication Union (ITU). **ITU-T Recommendation H.265: High Efficiency Video Coding, Audiovisual and Multimedia Systems**, Acessado em 22 jul. 2015. Online. Disponível em: <http://www.itu.int/rec/T-REC-H.265-201304-I>.

Sanchez, G, et-al. A complexity reduction algorithm for depth maps intra prediction on de 3D-HEVC, **Visual Communications and Image Processing Conference**, p.137-140, 2014

3D-HEVC-HTM. **Multimedia Communications with SVC, HEVC and SHVC**. Acessado em 22 jul. 2015. Online. Disponível em: http://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags