

HERANÇA DE PARÂMETROS DE MOVIMENTO NA CODIFICAÇÃO DE VÍDEOS 3D NO FORMATO MÚLTIPLAS VISTAS MAIS PROFUNDIDADE

MARCEL CORRÊA; BRUNO ZATT; MARCELO PORTO; LUCIANO AGOSTINI

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

1. INTRODUÇÃO

Avanços recentes em tecnologias relacionadas a vídeos 3D levam ao aumento crescente no interesse industrial e acadêmico por este tipo de vídeo, especialmente em alta definição. Como exemplo disso, o número de cinemas equipados com tecnologia para reprodução de filmes 3D, bem como o de filmes produzidos em 3D, aumentou drasticamente nos últimos anos. Não se limitando apenas a isso, a popularização de televisões, reprodutores *Blu-ray* e até mesmo dispositivos portáteis capazes de reproduzir conteúdo 3D levou esta tecnologia também às casas dos consumidores.

Os vídeos 3D proporcionam uma experiência visual com percepção de profundidade através do uso de *displays* especiais, capazes de projetar imagens levemente diferentes de maneira separada para cada olho. Estes *displays* incluem os do tipo estereoscópicos, que tipicamente reproduzem duas vistas gravadas simultaneamente com um par de câmeras, e que devem ser visualizadas com o auxílio de óculos especiais. Além destes, existem também os autoestereoscópicos, que não mostram apenas um par de vistas, mas sim um grande número que pode variar de 20 até mais de 50 vistas (MÜLLER, 2013). Este tipo de *display* permite que o usuário enxergue um par de vistas diferente, dependendo da posição dele em relação ao *display*, além de dispensar o uso de óculos especiais (BENZI, 2007).

O estado da arte para codificação de vídeos 3D atualmente é a extensão MVC (*Multiview Video Coding*) (VETRO, 2011) do padrão H.264/AVC. Porém, existem dois grandes problemas em relação ao MVC: Não permitir a codificação de canais de profundidade e apresentar um desempenho proibitivo conforme o número de vistas do vídeo aumenta.

Com a aprovação oficial no ano de 2013 do padrão HEVC (ITU-T, 2014), este sucedeu o H.264/AVC (WIEGAND, 2003) como o estado da arte em codificação de vídeos 2D. Com isso, esforços estão sendo feitos para a padronização de uma extensão do HEVC para codificação eficiente de vídeos 3D no formato múltiplas vistas mais profundidade (MVD - *Multiview plus Depth*). Neste formato, cada vista é associada a um canal de profundidade, que representa a distância dos objetos em relação à câmera. A Figura 1 ilustra um exemplo de um quadro de textura associado a um mapa de profundidade, onde no mapa os tons mais claros indicam maior proximidade em relação a câmera.

O formato MVD permite que a codificação de vídeo seja feita em um número bastante reduzido de vistas e seus respectivos mapas de profundidade, ao mesmo tempo em que permite que a decodificação gere vistas virtuais (extras) com base na informação codificada. Este processo de gerar vistas é chamado de síntese de vistas, e faz com que seja possível a visualização de um número de vistas muito maior que o número originalmente codificado.

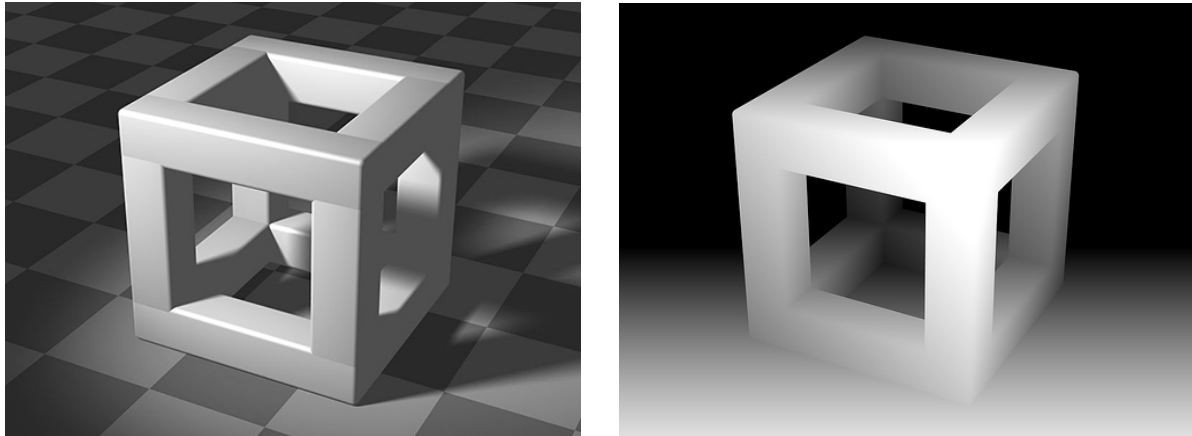


Figura 1 – Quadro de uma vista com informações de textura (esquerda) e mapa de profundidade associado (direita).

2. METODOLOGIA

No formato MVD, é preciso que os canais de profundidade também passem pelo processo de codificação, e como estas informações de profundidade apresentam características bastante específicas, novas técnicas estão sendo desenvolvidas pela comunidade para a codificação destes canais.

Este trabalho, ainda em estágio inicial, visa identificar quais técnicas apresentam ganhos mais significativos, e identificar também, possibilidades de aprimoramento destas técnicas. Este resumo aborda a técnica de herança de parâmetros de movimento (MPI – *Motion Parameter Inheritance*) (MÜLLER, 2013).

A ideia básica por trás do MPI está em utilizar os mesmos parâmetros de movimento, resultantes da codificação de um quadro de uma vista, na codificação do mapa de profundidade associado. Isso simplifica o processo custoso de codificação, uma vez que o esforço realizado para obtenção de partições e vetores de movimento ou disparidade poderá ser reaproveitado.

Esta técnica foi inserida da seguinte forma: A partição de textura co-localizada a partição atual de profundidade é tratada como a primeira candidata da etapa de *merge* do padrão HEVC. A utilização da sintaxe do *merge* é bastante conveniente, pois existem casos onde a herança de parâmetros permite que nenhum resíduo seja transmitido, e este é um comportamento similar ao do modo *skip*, que também é tratado na etapa de *merge*. O MPI, porém, apresenta também um custo, e este está na necessidade de verificar este candidato extra com a custosa avaliação de distorção (VSO – *View Synthesis Optimization*). O VSO irá avaliar qual candidato ao *merge* resultará em uma melhor qualidade ao final da codificação.

Por fim, o MPI apresenta também uma importante limitação. A técnica só é possível para uma dada partição de profundidade caso toda a partição co-localizada tenha sido codificada pela predição temporal ou de disparidade. Ainda, o vetor de movimento da partição co-localizada não pode apontar para uma partição que excede os limites do quadro.

3. RESULTADOS E DISCUSSÃO

De acordo com a extensiva avaliação realizada por WINKEN et al. (2012) com sequências e parâmetros de codificação levemente diferentes daqueles recomendados pelo documento das condições comuns de teste (JCT-3V, 2014), o MPI apresenta ganhos significativos.

Para o caso onde duas vistas foram codificadas, foi reportado a média de ganho de 0,05 dB PSNR na qualidade, e redução total de 1,4% no *bitrate*, o que significa uma redução de 18,15% no *bitrate* dos mapas de profundidade codificados, uma vez que o MPI não influencia na codificação das vistas de textura.

Já para o caso onde três vistas foram codificadas, foi observado um ganho de 0,04 dB PSNR na qualidade, e redução total de 1,12% no *bitrate*, que significa uma redução de 9,96% no *bitrate* dos mapas codificados.

Os ganhos pelo uso do MPI reportados por WINKEN et al. (2012) são de fato muito bons, mas não são contra intuitivos. Faz todo sentido aproveitar parâmetros de movimento de uma vista para seu mapa de profundidade, uma vez que ambos representam exatamente o mesmo cenário, do mesmo ponto de vista, no mesmo instante de tempo.

Com base no sucesso do MPI, pretende-se ao decorrer deste trabalho, realizar um refinamento desta técnica, ou até mesmo o desenvolvimento de uma técnica independente. Como ideia inicial, está sendo considerada a possibilidade de fazer a codificação dos mapas de profundidade antes das vistas, e herdar parâmetros no “caminho inverso” ao proposto pelo MPI. Esta ideia se baseia na característica dos mapas de profundidade de apresentarem bordas bem definidas que indicam bordas reais de objetos, e sendo assim, parâmetros de movimento obtidos na codificação inter-quadros ou inter-vistas destas bordas devem provavelmente ser adequados para herança na codificação das vistas.

4. CONCLUSÕES

Este trabalho descreveu a técnica de herança de parâmetros de movimento na codificação de vídeos 3D, no formato múltiplas vistas mais profundidade, utilizada no padrão emergente de codificação de vídeo 3D, o 3D-HEVC.

Esta técnica foi avaliada como responsável por ganhos significativos na codificação destes vídeos, e escolhida como ponto de partida para refinamentos que serão estudados neste trabalho de mestrado, com o intuito de aprimorar ainda mais a codificação de mapas de profundidade e/ou de vistas de textura em vídeos 3D.

5. REFERÊNCIAS BIBLIOGRÁFICAS

BENZIE, P.; et al. A Survey of 3DTV Displays: Techniques and Technologies. **IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY**, v.17, n.11, p.1647-1658, 2007.

ITU-T Recommendation Database: ITU-T H.265 (04/2013). Disponível em: <<http://www.itu.int/ITU-T/recommendations/rec.aspx?rec=11885>>. Acesso em 31 jul. 2014.

JCT-3V: Common Test Conditions of 3DV Core Experiments. Disponível em: <http://phenix.int-evry.fr/jct2/doc_end_user/documents/3_Geneva/wg11/JCT3V-C1100-v1.zip>. Acesso em 31 jul. 2014.

MÜLLER, 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.

VETRO, A.; WIEGAND, T.; SULLIVAN, G. Overview of the Stereo and Multiview Video Coding Extensions of the H.264/AVC Standard. **PROCEEDINGS OF THE IEEE, SPECIAL ISSUE 3D MEDIA DISPLAYS**, v.99, n.4, p.626-642, 2011.

WIEGAND, T.; SULLIVAN, G.; BJONTEGAARD, G.; LUTHRA, A. Overview of the H.264/AVC Video Coding Standard. **IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY**, v.13, n.7, p.560-576, 2003.

WINKEN, M.; SCHWARZ, H.; WIEGAND, T. Motion Vector Inheritance for High Efficiency 3D Video plus Depth Coding. In: **PICTURE CODING SYMPOSIUM**, Krakow, 2012. PCS Proceedings... Krakow: IEEE, 2012. p.53.