

HEURÍSTICA PARA REDUÇÃO DE COMPLEXIDADE NO CODIFICADOR DE MAPAS DE PROFUNDIDADE DO 3D-HEVC

RUHAN CONCEIÇÃO¹; GIOVANNI AVILA¹; GUILHERME CORRÊA¹;
MARCELO PORTO¹; BRUNO ZATT¹; LUCIANO AGOSTINI¹

¹Universidade Federal de Pelotas

{radconceicao,gdadaavila,gcorrea,porto,zatt,agostini}@inf.ufpel.edu.br

1. INTRODUÇÃO

Atualmente, presencia-se uma popularização gradual de dispositivos capazes de capturar, reproduzir e manipular vídeos digitais em três dimensões (3D), os quais viabilizam a percepção aprimorada – pelo espectador – da distância entre os objetos da cena e as câmeras utilizadas na gravação. Isto é possível apresentando-se, para cada cena, um par de vistas ao espectador, simulando o comportamento do sistema visual humano.

Dentre os *displays* utilizados para reproduzir vídeos 3D, destacam-se os estereoscópicos e auto estereoscópicos (TECH, 2015). O primeiro transmite ambas as vistas ao espectador, sendo que cada vista é filtrada para cada olho deste através do uso de óculos especiais para este fim. O segundo *display* descarta a necessidade por óculos que realizem a filtragem da imagem; neste caso, cada olho do espectador recebe uma dada vista de acordo com o ângulo de visão deste em relação à tela. Destaca-se que os *displays* auto estereoscópicos são capazes de transmitir uma elevada quantidade de vistas simultaneamente, e a posição do espectador em relação ao *display* define qual par de vistas este receberá (TECH, 2015).

Se por um lado sistemas auto estereoscópicos proporcionam uma experiência de entretenimento ainda mais elevada, por outro lado, os vídeos digitais utilizados nestes sistemas requerem uma enorme quantidade de dados para serem armazenados e/ou transmitidos. Sendo assim, faz-se necessário o uso de técnicas de codificação/compressão de vídeo, tal como de síntese de vistas nestes sistemas.

Basicamente, em um sistema auto estereoscópico que utilize síntese de vistas o transmissor processa um número reduzido de vistas espaçadas e as envia ao receptor. Este, por sua vez, decodifica o bitstream recebido, gerando então vistas intermediárias sintéticas através de alguma técnica de renderização de imagens baseada em profundidade (*Depth-Image-Based Rendering* – DIBR) (TECH, 2015). A geração de vistas sintéticas utilizando DIBR somente é possível utilizando-se mapas de profundidade, que são imagens em escala de cinza que representam a distância da câmera aos objetos da cena. Desta forma, percebe-se que o transmissor no sistema auto estereoscópico deve transmitir (após codificar) um mapa de profundidade correspondente para cada vista, o qual é capturado utilizando-se câmeras especiais para este fim.

O estado-da-arte em codificação de vídeos 2D é o padrão *High Efficiency Video Coding* (HEVC) (ITU-T, 2014), sendo sua extensão para vídeos 3D que utilizam mapas de profundidade o *3D-High Efficiency Video Coding* (3D-HEVC) (TECH, 2015). O uso do 3D-HEVC proporciona uma elevada compressão do vídeo 3D processado, contudo, o esforço computacional empregado no processo de codificação é extremamente elevado. Sendo assim, faz-se necessário o emprego de heurísticas que reduzam o tempo de codificação do 3D-HEVC, sem apresentar impactos significativos à eficiência de compressão do mesmo.

Tendo isto em vista, este trabalho propõe uma heurística de redução de complexidade (mensurada em redução de tempo de codificação) para o codificador de mapas de profundidade do 3D-HEVC. Tal solução consiste em testar primeiramente os modos *Skip* (ITU-T, 2014) e *Depth Intra Skip* (DIS) (CHEN, 2015), e somente avaliar os demais modos de codificação caso estes não apresentem um resultado satisfatório dado um critério apresentado a seguir. Desta forma, a heurística proposta visa evitar o teste de modos e ferramentas de codificação de mapas de profundidade potencialmente desnecessários.

2. METODOLOGIA

Primeiramente, avaliou-se a taxa de ocorrência dos modos *Skip* e DIS de acordo com o custo de codificação gerado para cada bloco do vídeo codificado. Destaca-se que o custo de codificação de um dado no modo/ferramenta no 3D-HEVC é dado utilizando a métrica custo taxa-distorção (*Rate-Distortion* – RD) (CHEN, 2015). Simulações utilizando o software de referência do 3D-HEVC, o *3D-HEVC Test Model* (3D-HTM) versão 16.0 (CHEN, 2015), demonstraram que, quanto maior custo RD gerado pelos modos *Skip* e DIS, menor a taxa de ocorrência destes modos. Sendo assim, a probabilidade destes modos ocorrerem é inversamente proporcional ao custo RD gerado. A Figura 1 ilustra um gráfico que demonstra a taxa de ocorrência dos modos (a) *Skip* e (b) DIS em função do custo RD gerado. Os dados utilizados na construção do gráfico foram obtidos a partir das simulações realizadas.

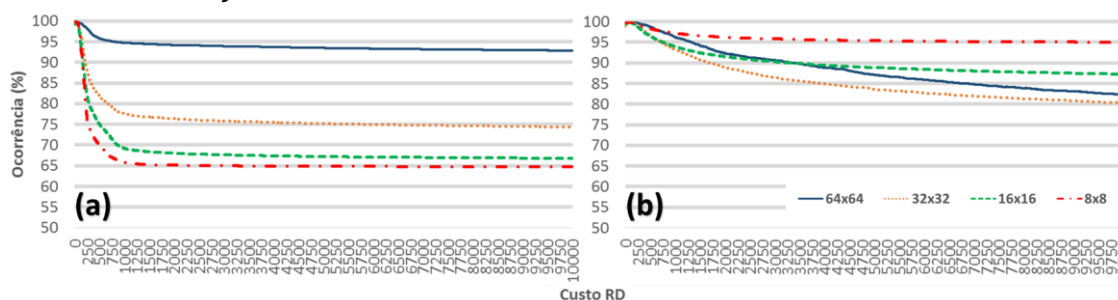


Figura 1. Ocorrência dos modos (a) *Skip* e (b) DIS em função do custo RD gerado

Uma vez identificado que a probabilidade de ocorrência dos modos *Skip* e DIS é inversamente proporcional ao custo RD gerado por estes, desenvolveu-se a heurística foco deste trabalho.

O funcionamento da heurística é dado da seguinte forma: primeiramente, os oito primeiros conjuntos de vistas do vídeo 3D é codificado – utilizando o 3D-HEVC sem nenhuma otimização –, sendo então coletado dados de taxa de ocorrência dos modos *Skip* e DIS em função do custo RD gerado por estes. A seguir, com base nos dados obtidos, são estipulados valores limiares de custo RD que definirão se os demais modos de codificação serão, ou não, testados após o processamento dos modos *Skip* e DIS.

Os valores limiares utilizados na heurística são definidos de acordo com uma taxa de ocorrência alvo de $X\%$. Sendo assim, caso deseje-se que a heurística tenha uma precisão de $X\%$, o valor limiar r escolhido pela heurística será o custo RD r que resulta em uma taxa de ocorrência de $X\%$ para o modo *Skip*/DIS.

Destaca-se que dois valores limiares são utilizados nesta heurística: TH1 e TH2. O primeiro refere-se ao maior custo RD resultado pelo *Skip* no qual a heurística evitará o teste dos demais modos de codificação (inclusive o DIS). O segundo apresenta o maior custo RD que um dos modos *Skip* e DIS poderá conter para que o teste das demais ferramentas de codificação seja evitado. É

importante destacar que a heurística proposta prioriza a seleção do modo *Skip* em relação ao modo *DIS* devido aos ganhos de codificação superiores proporcionados por aquela ferramenta em relação a esta. A Figura 2 apresenta um fluxograma do algoritmo empregado na aplicação da heurística proposta neste artigo.

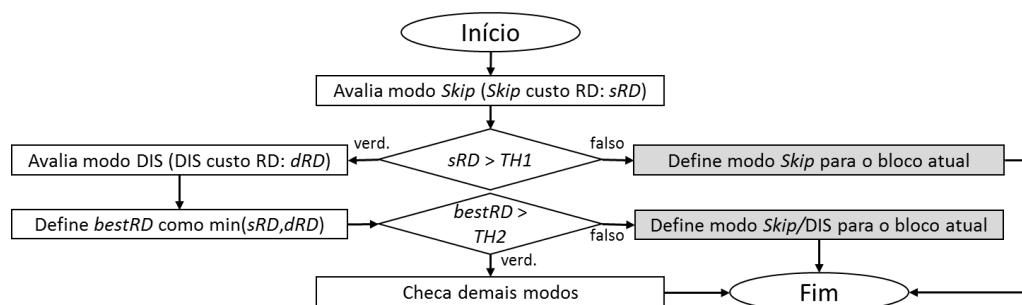


Figura 2. Fluxograma da heurística proposta

3. RESULTADOS E DISCUSSÃO

A heurística foi avaliada utilizando o 3D-HTM versão 16.0 de acordo com as condições comuns de teste (CCT) (RUSANOVSKYY, 2013). Nos experimentos realizados, definiu-se em 99% a taxa de ocorrência alvo para a estipulação dos valores limiares de custo RD. Em outras palavras, estes valores limiares foram definidos dinamicamente pela heurística (para cada experimento) visando proporcionar uma precisão 99% na tomada de decisão de modo antecipada.

Há na literatura alguns trabalhos que abordam estratégias de redução de tempo de codificação de mapas de profundidade do 3D-HEVC, tais como (SANCHEZ, 2014), (PARK, 2016) e (ZHANG, 2015). Um comparativo entre os resultados de simulação obtidos neste trabalho com os relacionados é apresentado na Tabela 1. Os resultados avaliados na comparação são: impacto à eficiência de codificação da textura e das vistas sintetizadas, a qual é avaliada em termos de *Bjontegaard delta bitrate* (BD-BR) (BJONTEGAARD, 2001); redução (Red.) de tempo de codificação total do 3D-HEVC; e redução do tempo de codificação dos mapas de profundidade apenas. Estes resultados são obtidos através dos experimentos realizados, sendo calculados de acordo com as CCT.

Tabela 1. Comparação com trabalhos relacionados

Trabalho	Textura (BD-BR)	Síntese (BD-BR)	Red. Tempo Total	Red. Tempo Profundidade
(SANCHEZ, 2014)	0,027%	-0,064%	5,9%	23,8%
(PARK, 2016)	-0,030%	0,130%	-	22,2%
(ZHANG, 2015)	-	1,030%	-	27,9%
Este Trabalho	0,012%	0,067%	9,6%	24,4%

Como é possível perceber na Tabela 1, a heurística proposta é capaz de reduzir – em média – 24,4% o tempo de codificação dos mapas de profundidade no 3D-HEVC, sendo a segunda melhor solução neste quesito quando comparado aos trabalhos relacionados. Entretanto, é importante destacar que solução de (ZHANG, 2015) impacta em 1,030% a eficiência de codificação das vistas sintetizadas, contra apenas 0,067% da heurística proposta. Isto, somado aos dados apresentados na Tabela 1, demonstra a relevância da solução proposta neste artigo.

4. CONCLUSÕES

Este trabalho apresentou uma heurística para a redução da complexidade – mensurada em termos de tempo de codificação – do codificador de mapas de profundidade do padrão 3D-HEVC. Tal solução visa avaliar primeiramente os modos *Skip* e *DIS* e, de acordo com o custo RD gerado por estes modos, a avaliação dos demais modos de codificação evitada. Resultados experimentais demonstraram que a heurística é capaz de reduzir em média 24,4% o tempo total de codificação dos mapas de profundidade. Comparado com trabalhos relacionados da área de codificação de vídeo, os quais foram publicados nas principais revistas e eventos internacionais da área, conclui-se que a heurística proposta neste artigo apresenta grande relevância científica. Como trabalho futuro, pretende-se expandir esta solução ao codificador de textura do 3D-HEVC, visando reduzir ainda mais o tempo total de codificação do 3D-HEVC.

5. REFERÊNCIAS BIBLIOGRÁFICAS

BJONTEGAARD, G. **Calculation of average PSNR differences between RDcurves**. [S.l.]: ITU-T, 2001.

CHEN, Y.; TECH, G.; WEGNER, K.; YEA, S. **Test Model 11 of 3D-HEVC and MVHEVC**. [S.l.]: Joint Collaborative Team on 3D Video Coding Extension, 2015.

ITU-T. **High efficiency video coding**: Recommendation ITU-T H.265. [S.l.]: International Telecommunication Union: Telecommunication Standardization Sector of ITU, 2014.

PARK, C.-S. Edge-Based Intramode Selection for Depth-Map Coding in 3D-HEVC. **Image Processing, IEEE Transactions on**, [S.l.], v.24, n.1, p.155–162, Jan 2015.

RUSANOVSKYY, D.; MULLER, K.; VETRO, A. **Common Test Conditions of 3DV Core Experiments**. [S.l.]: Joint Collaborative Team on 3D Video Coding Extension, 2013.

SANCHEZ, G.; SALDANHA, M.; BALOTA, G.; ZATT, B.; PORTO, M.; AGOSTINI, L. Complexity reduction for 3D-HEVC depth maps intra-frame prediction using simplified edge detector algorithm. In: **IMAGE PROCESSING (ICIP), 2014 IEEE INTERNATIONAL CONFERENCE ON**, 2014. **Anais. . .** [S.l.: s.n.], 2014. p.3209–3213.

TECH, G.; CHEN, Y.; MULLER, K.; OHM, J.-R.; VETRO, A.; WANG, Y.-K. Overview of the Multiview and 3D Extensions of High Efficiency Video Coding. **Circuits and Systems for Video Technology, IEEE Transactions on**, [S.l.], v.PP, n.99, p.1–1, 2015.

ZHANG, N.; CHEN, Y.-W.; LIN, J.-L.; FAN, X.; MA, S.; ZHAO, D.; GAO, W. Improved disparity vector derivation in 3D-HEVC. In: **VISUAL COMMUNICATIONS AND IMAGE PROCESSING (VCIP), 2013, 2013. Anais. . .** [S.l.: s.n.], 2013. p.1–5.