

## REDUÇÃO DE COMPLEXIDADE NA TRANSCODIFICAÇÃO HOMOGÊNEA NO PADRÃO HEVC PARA ADAPTAÇÃO DE TAXA DE BITS

THIAGO LUIZ ALVES BUBOLZ; RUHAN CONCEIÇÃO; MARCELO PORTO;  
GUILHERME CORRÉA

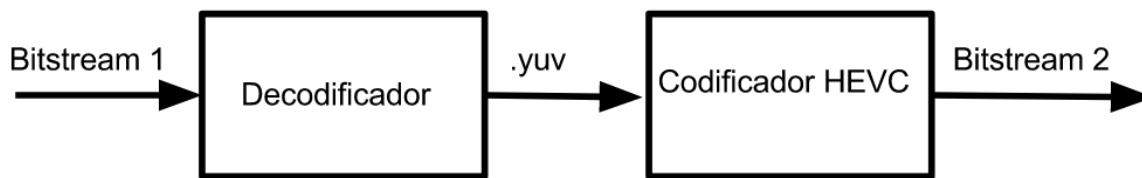
Universidade Federal de Pelotas – Video Technology Research Group (ViTech)  
*{tlabubolz, radconceicao, porto, gcorrea}@inf.ufpel.edu.br*

### 1. INTRODUÇÃO

O *High Efficiency Video Coding* (HEVC) (SULLIVAN, 2012) é o atual estado da arte entre os padrões de codificação de vídeo. Entre as principais metas em destaque prometidas e atingidas por este padrão durante seu desenvolvimento, destaca-se atingir o dobro de compressão em relação ao seu antecessor, o padrão H.264/AVC (WIEGAND, 2003), mantendo a mesma qualidade de imagem e complexidade computacional. Embora tenha conseguido grandes taxas de compressão e manter a qualidade da imagem, o HEVC não conseguiu manter a complexidade computacional do seu antecessor. Alguns trabalhos publicados mostram que o HEVC atingiu um esforço computacional até 500% maior que o seu antecessor H.264/AVC (CORREA, 2012).

A transcodificação de vídeo pode ser utilizada com o intuito de modificar o padrão em que o vídeo está codificado, no caso uma transcodificação heterogênea. Alternativamente, a transcodificação pode manter o vídeo no mesmo padrão e fazer alguma alteração em suas informações, como a adição de uma marca d'água ou a diminuição da taxa de bits em uma compressão maior.

Segundo um estudo da CISCO (2017), o uso de serviços de *streaming* e download de conteúdo de vídeo ultrapassarão os 80% de uso total da banda mundial até o ano de 2021, tornando o uso destes serviços um tema bastante importante quando se fala em compressão de mídia. Devido ao alto uso de serviços de *streaming* de vídeo como *Youtube* e *Netflix*, a transcodificação se transformou em uma tarefa importante para este tipo de serviço, já que é necessário manter várias versões de um mesmo vídeo com taxas de bits diferentes. Além disso, pode a transcodificação homogênea pode ser realizada em tempo real no lado servidor, de forma a adequar o conteúdo à largura de banda dos seus clientes.

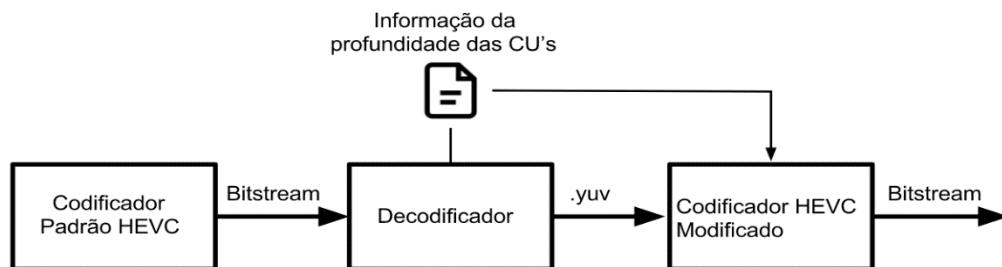


**Figura 1.** Transcodificador simples

A Figura 1 mostra como funciona a transcodificação homogênea de forma simples e de onde são retiradas as informações que são usadas na segunda codificação do vídeo para as comparações das profundidades.

Este trabalho propõe uma diminuição na complexidade computacional na transcodificação homogênea de vídeo no padrão HEVC, para isso foi feito um estudo estatístico com as profundidades de partições das *Coding Units* (CU's)

geradas na codificação para diferentes porcentagens de taxa de bits, cada frame do vídeo é dividido em multiplas CU's que são estruturas de *quad-trees* que suportam divisões de 64x64 que podem ser subdivididas em 32x32, novamente subdivididas em 16x16 e por fim chegam até divisões de 8x8, assim foi escolhido um limiar máximo para limitar a busca na recursão dessa profundidade em codificações seguintes.



**Figura 2.** Esquema de Funcionamento de um Transcodificador Homogêneo

## 2. METODOLOGIA

Esta seção aborda a metodologia utilizada para reproduzir os testes propostos para este trabalho. Inicialmente, foi necessária uma avaliação estatística sobre uma série de vídeos de diferentes resoluções e qualidades. Para isso, foi comparada a profundidade em que uma CU foi particionada em uma primeira codificação com a profundidade da mesma CU recodificada em outras taxas de bits. A média dos dados de todos os vídeos estão expostos na Tabela 1. Foi observado que boa parte das CUs se particionam em um tamanho igual ao utilizado na codificação original, ou aumentam o tamanho. Em uma proporção bem menor, ocorre de CUs se particionarem em tamanhos menores em uma codificação com taxas de bits menores.

**Tabela 1.** Média das Porcentagens das Partições das CUs

Particionamento da CU	64x64	32x32	16x16	8x8
<b>64x64</b>	<b>80,22%</b>	17,61%	2,03%	0,13%
<b>32x32</b>	37,01%	<b>55,95%</b>	6,57%	0,47%
<b>16x16</b>	22,13%	37,83%	<b>37,69%</b>	2,34%
<b>8x8</b>	8,95%	21,72%	22,94%	<b>46,38%</b>

Com base nesses dados, para diminuir o tempo de codificação de vídeo, foi escolhido um limite avaliado na média das porcentagens para que o codificador não precisasse testar particionamentos menores que o necessário. Portanto, todos os valores que fossem abaixo de 8% foram retirados da busca. Na Tabela 1, as divisões em verde entram na busca do codificador e os elementos em vermelho foram retirados, os vídeos testados nesta etapa foram *BlowingBubbles*, *RaceHorses*, *SlideEditing*, *KristenAndSara*, *BasketBallDrive* e *Traffic*. Portanto, no processo de transcodificação, a abordagem compara a profundidade atual com a profundidade em que o codificador particionou previamente a mesma CU durante a codificação original. Caso este valor chegue no limite definido, a busca para a nova profundidade é escolhida entre as opções previamente calculadas. Caso uma CU tenha sido particionada em 64x64, na transcodificação a busca pararia na profundidade 32x32, se tivesse sido particionada em 32x32 buscaria também

até 32x32, para particionamentos 16x16 e 8x8 a busca para respectivamente em 16x6 e 8x8.

Após as análises foram feitos testes para um vídeo de cada classe da *Common Test Conditions* (CTC). Os vídeos utilizados foram *ToddlerFountain*, *RollerCoaster*, *BQTerrace*, *PartyScene*, *BQSquare*, *Johnny* e *SlideShow*. Todos os vídeos foram primeiramente codificados com um *Quantization Parameter* (QP) 22. Após, foram calculadas suas respectivas taxas de bits para 80%, 60%, 40% e 20% da taxa original e então os vídeos foram transcodificados para essas taxas com o codificador HEVC modificado conforme explicado neste trabalho. Após todas as transcodificações terminadas, os dados foram recolhidos para uma análise objetiva de qualidade, que será abordada na próxima seção.

### 3. RESULTADOS E DISCUSSÃO

Esta seção discute os dados obtidos através de uma análise objetiva dos vídeos transcodificados. Analisando os dados de todas as transcodificações, foram calculados os valores de BD-rate (isto é, a eficiência da compressão) e a redução percentual de tempo médio das transcodificações realizadas conforme a abordagem proposta neste trabalho. A Tabela 2 mostra estes resultados para cada um dos vídeos testados. Os vídeos são de diferentes classes dentro da CTC e de diferentes resoluções. O vídeo que conseguiu uma maior redução percentual média de tempo de transcodificação foi o *RollerCoaster*, porém este vídeo também apresentou o segundo maior BD-rate, ou seja, a segunda pior eficiência de compressão. Já o vídeo com o menor BD-rate (isto é, melhor eficiência de compressão) é o vídeo *BQSquare*, que por sua vez foi aquele que obteve a menor redução percentual média de tempo.

Para uma garantia de corretude dos resultados, todos os testes foram realizados em uma mesma máquina com as mesmas condições. Neste caso, a diferença no tempo de transcodificação do modelo proposto em relação ao HEVC é diretamente proporcional à profundidade em que as CUs foram codificadas na primeira etapa. Portanto, quanto maior o número de CUs de maior profundidade, como por exemplo CUs particionadas em 8x8, menor será o tempo reduzido, pois para uma CU particionada previamente em 8x8 o transcodificador irá permitir uma busca até a mesma profundidade.

**Tabela 2:** Resultados de Redução Percentual de Tempo Médio e BD-rate

CLASSE	VÍDEOS		
		TEMPO	BD-rate
A1	<b>ToddlerFountain_4096x2160</b>	39,20%	0,4352
A2	<b>RollerCoaster_4096x2160</b>	56,37%	1,1020
B	<b>BQTerrace_1920x1080</b>	32,40%	0,6487
C	<b>PartyScene_832x480</b>	31,61%	0,5160
D	<b>BQSquare_416x240</b>	31,41%	0,3712
E	<b>Johnny_1280x720</b>	56,34%	0,7955
F	<b>SlideShow_1280x720</b>	53,22%	1,5797
<b>Media</b>		42,93%	0,7783

A média de redução percentual no tempo entre os vídeos testados foi de 42,93%. Já a média de BD-rate (perda de eficiência de compressão) foi de apenas 0,7783%. Isso mostra uma grande redução no tempo de codificação de vídeo com uma perda pequena na taxa de compressão ou na qualidade do vídeo

codificado. Os vídeos com as menores resoluções, como *PartyScene*, com resolução 832x480, e *BQSquare*, com resolução 416x240, mostraram bons números de BD-rate, mas uma diminuição pequena do tempo de codificação. Isso pode ocorrer já que um quadro de resolução 416x240 possui 25 CTUs, o que torna mais difícil a sua divisão em CUs pequenas devido à heterogeneidade das informações em cada CU.

Com uma grande redução no tempo de transcodificação e uma pequena queda na eficiência de codificação, a abordagem utilizada se mostrou bastante favorável para diferentes tipos de vídeos e em diferentes resoluções. Apesar de uma grande redução no tempo, esta abordagem por si só não é capaz de possibilitar transcodificação em tempo real, que seria fundamental para sistemas de *streaming* que necessitam entregar aos seus clientes vídeos em diferentes taxas de bits, porém se mostra bastante eficaz caso o sistema deseje manter todas as versões armazenadas em servidor. Caso a abordagem proposta seja aliada a outras técnicas de redução de complexidade, uma diminuição ainda maior do tempo de transcodificação poderá ser atingido, chegando mais próximo de uma transcodificação em tempo real.

#### 4. CONCLUSÕES

A abordagem apresentada neste resumo consegue cumprir a função proposta de diminuir o custo computacional para a transcodificação de vídeo e consegue manter a eficiência de codificação relativamente estável, como apresentado na avaliação objetiva. A abordagem para a redução da complexidade é simples e baseia-se em análises de dados estatísticos de divisão de CUs na primeira codificação para a tomada de decisão sobre quais profundidades devem ser testadas na recodificação do vídeo.

Para trabalhos futuros, serão comparados os resultados com mais vídeos para uma obtenção de resultados mais concretos. Aliado a este trabalho, pretende-se realizar a implementação de outras técnicas que possam auxiliar ainda mais na redução da complexidade, como a implementação da mesma abordagem para as decisões de *Prediction Units* (PUs) e *Transform Units* (TUs) no HEVC. Também é intenção adicionar alguma estratégia de aprendizado de máquina para que as decisões de busca na quad-tree sejam tomadas de uma maneira mais adequada do que uma limitação fixa, como a apresentada neste trabalho.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

- SULLIVAN G. J. et al. 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.
- WIEGAND, T. et al. 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.
- CORREA, G. et al. Overview of the H.264/AVC video coding standard. **IEEE Transactions on Circuits and Systems for Video Technology**, v.22, p. 1899-1909, 2012.
- Cisco Intersight. **Cisco Visual Networking Index: Forecast and Methodology, 2016–2021**. Cisco Website, 15 de Setembro, 2017. Acessado em 26 Set. 2017. Online. Disponível em: <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html>