

ESTUDO E IMPLEMENTAÇÃO DE DESCRIPTORES DE IMAGENS PARA REDUÇÃO DE COMPLEXIDADE COMPUTACIONAL DO PROCESSO DE PARTICIONAMENTO DE QUADROS EM CODIFICADORES HEVC

ELIÉZER RIBEIRO; GUILHERME CORRÉA; LUCIANO AGOSTINI;
BRUNO ZATT

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

1. INTRODUÇÃO

Atualmente, há uma crescente demanda por vídeos de alta resolução, pois os consumidores de conteúdo multimídia desejam cada vez mais uma melhor qualidade de imagem. Entretanto, este tipo de vídeo é representado por uma grande quantidade de dados, exigindo um enorme espaço de armazenamento e causando problemas sérios em termos de transmissão e recepção. Para resolver este problema, são utilizados algoritmos para compressão desses vídeos, possibilitando o armazenamento e transmissão do mesmo vídeo com uma quantidade de bits bem menor do que a original e com perdas de qualidade muitas vezes imperceptíveis ao olho humano.

O processo de codificação de vídeos é dado por uma variedade de ferramentas e técnicas, que foram e ainda vem sendo desenvolvidas, tornando a pesquisa nesta área muito ativa. Atualmente, o padrão de codificação de vídeo mais eficiente disponível é o *High Efficiency Video Coding* (HEVC) (JCT-VC, 2013). O HEVC atinge taxas de compressão muito maiores que os seus predecessores, porém apresenta uma complexidade computacional enorme, o que é um fator relevante a se observar em codificadores, especialmente quando consideramos os vídeos de alta resolução, que requerem uma quantidade elevada de processamento.

Uma das principais inovações do padrão HEVC é a utilização de estruturas de particionamento de quadro muito flexíveis: as *Coding Units* (CU), as *Prediction Units* (PU) e as *Transform Units* (TU). A CU é a uma unidade básica que permite a divisão recursiva de um bloco da imagem em quatro blocos de tamanhos iguais, sob a forma de árvore quadrática. Todos os demais particionamentos são realizadas a partir das CUs, de forma que esta é a estrutura mais importante do ponto de vista da complexidade computacional do codificador HEVC. A Figura 1 mostra quatro regiões de 64x64 pixels da imagem divididas em diversas CUs sob a forma de árvore quadrática (CORREA, 2015). Para definir a melhor estrutura de árvore quadrática e quais os sub-particionamentos devem ocorrer, o codificador HEVC de referência testa todas as possibilidades de árvore (i.e., todos os possíveis sub-particionamentos de CU) e escolhe a combinação de CUs que resulta na melhor eficiência de compressão. Entretanto, esta decisão envolve uma enorme quantidade de cálculos, pois todo o processo de codificação para as diversas possibilidades de árvores quadráticas deve ser executado, enquanto que apenas uma das possibilidades é escolhida no final. Logo, para reduzir a complexidade computacional do codificador HEVC, é essencial que encontrem-se formas de simplificar o processo de decisão de particionamento de CUs.

Este trabalho tem como objetivo apresentar uma análise estatística sobre características das porções das imagens representadas nas CUs, correlacionando tais características com as decisões de sub-particionamento realizadas pelo

codificador HEVC. Tal análise tem por objetivo a identificação de características que possam ser utilizadas futuramente no desenvolvimento de uma heurística ou modelo que realize decisões rápidas de sub-particionamento de CU, evitando os testes exaustivos realizados no codificador HEVC de referência.

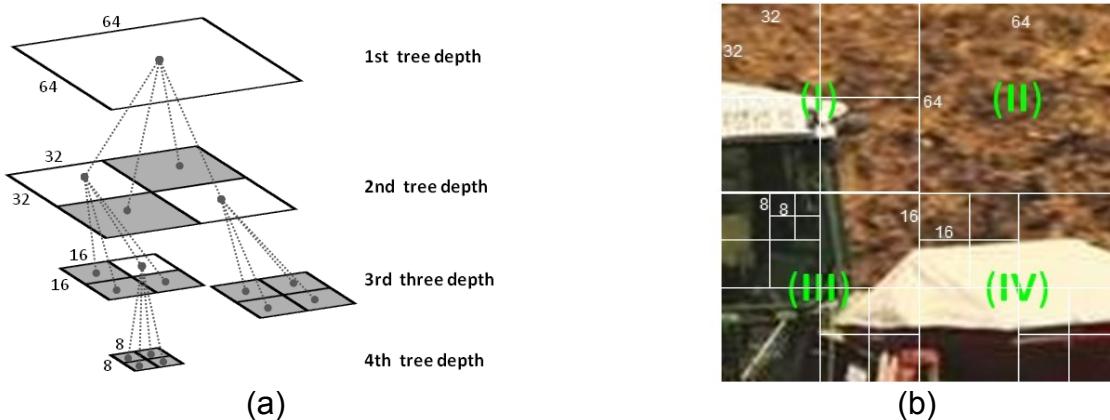


Figura 1: Exemplo de quatro árvores quadráticas (a) e suas respectivas CUs (b).

2. METODOLOGIA

Uma característica interessante dos vídeos de altas resoluções é que existe uma tendência maior de haver mais regiões com textura homogênea em comparação com os vídeos de baixa resolução. Áreas mais homogêneas são geralmente codificadas com CUs maiores, enquanto que áreas heterogêneas são codificadas com CUs menores. Os dados analisados neste trabalho descrevem características das CUs relacionadas à sua textura, utilizando as amostras de luminância do vídeo original para os cálculos. Todos os descritores foram implementados em C e o programa desenvolvido foi utilizado para coletar dados dos vídeos *BasketballDrive*, *BQTerrace*, *Cactus* e *Kimono1*, todos de resolução 1080p (1920x1080). Foram analisados dados obtidos a partir de amostras de luminância de CUs 64x64, que correspondem à maior dimensão que as CUs podem assumir nos vídeos codificados conforme o padrão HEVC. Os parágrafos a seguir apresentam todos os descritores implementados neste trabalho.

A Homogeneidade da CU foi calculada através de um filtro detector de bordas, que consiste em uma matriz 3x3 que procura detectar, ponto a ponto, as bordas presentes em uma imagem. A matriz está apresentada na (Eq. 1). Note que a soma algébrica dos seus pesos é igual a zero, de forma que, quando aplicado a pontos em regiões totalmente homogêneas de uma imagem, o filtro resultará em um valor igual a zero ou em valores muito pequenos. Os valores obtidos ponto a ponto na aplicação do filtro são comparados com um *threshold* positivo próximo de zero. O número de pontos com resultado do filtro menor que o *threshold* define o nível de homogeneidade da CU. Quanto mais pontos com resultado próximo a zero, mais homogênea é a CU.

$$X = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \times A \quad (\text{Eq. 1})$$

O Gradiente, também chamado de coeficiente de assimetria, foi calculado através de um filtro de detecção de bordas horizontais e verticais sobre as amostras de luminância das CUs 64x64. O filtro é composto por duas matrizes 3x3 apresentadas em (Eq. 2) e (Eq. 3), que são convoluídas com a imagem

original para calcular aproximação das derivadas – uma para as variações horizontais e outra para as verticais. Sendo A a imagem original, então Gx e Gy serão duas imagens que, em cada ponto, contêm uma aproximação às derivadas horizontais e verticais de A . Portanto, a magnitude do gradiente G é dada pela (Eq. 4).

$$Gx = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \times A \quad (\text{Eq. 2})$$

$$Gy = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \times A \quad (\text{Eq. 3})$$

$$G = \sqrt{Gx^2 + Gy^2} \quad (\text{Eq. 4})$$

A entropia, ou grau de incerteza, é um valor diretamente proporcional à despadronização de uma imagem (neste caso, de uma CU). A entropia é calculada através da (Eq. 5), onde M é numero total de tons diferentes presentes na imagem e p_i é a probabilidade de ocorrência do i -ésimo tom.

$$E_x = \sum_{i=0}^{M-1} (p_i \log_2(p_i)) \quad (\text{Eq. 5})$$

O último descritor calculado foi a variância da imagem, que representa o valor de desvio dos níveis de cinza da imagem em relação ao nível de cinza médio, onde o valor médio (μ_i) de uma CU i é calculado pelo somatório dos valores de luminância de todos os pixels da CU, dividido pelo número total de pixels na CU (n). A variância da CU representa o valor de desvio dos níveis de cinza da imagem (nc) em relação ao nível de cinza médio, conforme representado na (Eq. 6).

$$\partial_i^2 = \frac{\sum_{k=0}^n (nc_{ki} - \mu_i)^2}{n - 1} \quad (\text{Eq. 6})$$

Além de todos os descritores apresentados nos parágrafos anteriores, também foi salva, para cada CU 64x64, a informação sobre o particionamento da CU 64x64 em sub-CUs 32x32 (i.e., *Split* ou *Não Split*). Os descritores e tal informação foram utilizados na análise estatística cujos resultados estão apresentados na próxima sessão.

3. RESULTADOS E DISCUSSÃO

A Tabela 1 e a Tabela 2 apresentam a média dos atributos coletados no processo descrito na seção 2 para QPs 22, 27, 32 e 37, agrupados em dois conjuntos: aquele referente às CUs que sofreram sub-particionamento no processo de codificação HEVC e aquele referente às CUs que não sofreram sub-particionamento. Além disso, a tabela apresenta valores em termos de ganho de informação (GI), uma métrica amplamente utilizada por algoritmos de *data mining* e *machine learning* para medir a relevância de utilização de uma variável num processo de classificação. O GI equivale à diferença entre o número de bits necessários para transmitir a identidade de um dado antes e após a sua classificação com base numa variável em questão. Quanto maior o GI de uma determinada variável, maior é a sua relevância em um processo de classificação.

A Tabela 1 e Tabela 2 mostram que há uma diferença significativa entre as médias de alguns atributos coletados nos dois casos analisados (*Split* e *Não Split*), de forma que essas informações podem ser utilizadas em uma futura heurística com foco em redução de complexidade para o codificador HEVC. As variáveis que apresentaram as maiores diferenças entre os dois conjuntos foram o gradiente, a homogeneidade e a média. Estas também foram as variáveis que apresentaram maiores valores de GI.

Tabela 1: Resultados para os atributos analisados, QP 22 e QP 27.

Atributo	QP 22			QP 27		
	Split	Não Split	GI	Split	Não Split	GI
Gradiente	136,677	73,818	0,12	135,966	99,267	0,05
Homogeneidade	140	26	0,10	138	73	0,03
Média	447,368	499,120	0,06	430,975	492,794	0,06
Borda Horiz.	13,994	15,592	0,05	13,485	15,399	0,05
Borda Vert.	13,907	15,542	0,05	13,407	15,326	0,05
Entropia	1,582	1,708	0,07	1,578	1,662	0,05
Variância	9.8×10^{14}	1.2×10^{15}	0,06	9×10^{14}	1.2×10^{15}	0,06

Tabela 2: Resultados para os atributos analisados, QP 22 e QP 27.

Atributo	QP 32			QP 37		
	Split	Não Split	GI	Split	Não Split	GI
Gradiente	125,692	111,869	0,04	125,711	70,637	0,04
Homogeneidade	111	99	0,01	107	46	0,01
Média	412,695	488,081	0,07	412,250	528,587	0,06
Borda Horiz.	12,922	15,252	0,06	12,909	16,537	0,05
Borda Vert.	12,831	15,184	0,06	12,824	16,352	0,05
Entropia	1,585	1,640	0,05	1,577	1,732	0,05
Variância	8.2×10^{14}	1.2×10^{15}	0,06	8.2×10^{14}	1.2×10^{15}	0,06

4. CONCLUSÕES

Este trabalho apresentou uma análise estatística de alguns atributos descritores de imagens, com relação à informação de sub-particionamento de CU 64x64 no codificador HEVC. O trabalho apresentou análises iniciais sobre algumas variáveis e demonstrou a existência de correlação entre algumas características dos vídeos e as decisões de sub-particionamento de CUs.

Em trabalhos futuros, pretende-se aprofundar a análise realizada, incluindo a investigação e implementação de outros atributos, e desenvolver técnicas de redução de complexidade computacional do codificador HEVC através da sua utilização.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- JCT-VC. High Efficiency Video Coding text specification draft 10, doc. JCTVC-L1003, Genebra, Suíça, 2013.
- CORRÊA, G.; ASSUNÇÃO, P. A.; AGOSTINI, L.; CRUZ, L. A.. Fast HEVC Encoding Decisions Using Data Mining , **IEEE Trans. on Circuits and Systems for Video Tech.**, Vol. 25 , N. 4 , pp. 660 - 673 , Abril, 2015.