

ANÁLISE DE REDUÇÃO DO VOLUME DE DADOS PROCESSADOS PELO MÓDULO DAS TRANSFORMADAS DO CODIFICADOR DE VÍDEO HEVC

Jones William Göebel; José Cláudio de Souza Jr.;
Luciano Agostini; Marcelo Porto; Bruno Zatt;

*Universidade Federal de Pelotas – Grupo de Arquitetura e circuito Integrado –
{ jwgoebel, jcdsouza, agostini, msporto, brunozatt }@inf.ufpel.edu.br*

1. INTRODUÇÃO

Atualmente, a qualidade dos vídeos digitais têm sido ampliada de forma contínua. Além disto, os vídeos são cada vez mais utilizados em diversos dispositivos, como smartphones, televisores digitais, blu-ray players, etc. Por isso, conveniu-se o a utilização de codificadores e decodificadores de vídeo, de forma que os dispositivos sejam capazes de transmitir vídeos em tempo real e armazená-los. Com este cenário, taxa de compressão, qualidade de vídeo, tempo de processamento e consumo de energia são questões relevantes.

Um dos maiores problemas na codificação de vídeos é que o aumento das resoluções de vídeos implica em processar mais informações, o que se torna uma tarefa difícil para processadores comuns. Para solucionar este problema, uma das alternativas é a implementação de arquiteturas de *hardware* dedicadas para codificação de vídeo. O estado da arte em codificação de vídeo, atualmente, é o padrão HEVC (JCT-VC, 2013), que tem como objetivo atingir uma taxa de 50% mais compressão em relação ao seu antecessor, mantendo a mesma qualidade do vídeo. Para que o HEVC atingisse essa taxa de compressão, houve um aumento de complexidade computacional (Bossen, 2012) em relação ao seu antecessor, o H.264/AVC (ITU, 2005).

O princípio básico da compressão de vídeos digitais consiste na exploração das informações redundantes. As redundâncias que podem ser encontradas nos vídeos são espacial, temporal ou entrópica. A redundância espacial ocorre pela similaridade de uma amostra com seus vizinhos dentro um mesmo quadro do vídeo. A redundância temporal é formada a partir de informações replicadas entre os quadros que compõe o vídeo. Por fim, a redundância entrópica está relacionada com a quantidade de informação representada por cada símbolo do vídeo. A partir da diferença entre o quadro que está sendo codificado, e os quadros de referência, são gerados os resíduos. Os resíduos são manipulados pelas etapas chamadas de transformada, quantização e entropia.

O foco deste trabalho está na transformada (T), onde as amostras são transformadas do domínio espacial para o domínio das frequências. Para realizar esta mudança de domínios, o HEVC utiliza a transformada discreta dos cossenos (DCT) ou a transformada discreta dos senos (DST). Dentro do processo realizado pelo codificador, é necessário montar o quadro reconstituído que será utilizado como quadro de referência pela etapa de predição. Desta forma é necessário realizar parte do processo de decodificação dentro do codificador, e sendo assim, necessitamos do processo de transformação inverso, realizado pelas transformadas inversas (TI).

Um problema encontrado nas transformadas é que, na etapa de codificação, todos os tamanhos definidos pelo padrão são processados. O objetivo deste trabalho é apresentar uma avaliação para reduzir o número de blocos transformados no processo de codificação de vídeos em resoluções

HD1080p (1920x1080 pixels), sem comprometer de maneira significativa a qualidade visual dos vídeos codificados.

2. METODOLOGIA

A DCT-2D genérica, está definida em (1) e (2), onde N e M são o número de pontos da DCT, $F(u, v)$ é a entrada para a posição (u, v) da matriz (bloco) de entrada, e $f(i, j)$ é o coeficiente de saída pertencente a matriz de saída. Uma inovação que foi inserida no HEVC é a estrutura de dados utilizada para a transformada, chamada de unidade de transformação (TU). Outra inovação foi a utilização de 4 tamanhos diferentes para os blocos das transformadas, sendo eles sempre quadrados: 32x32, 16x16, 8x8 e 4x4 (Sullivan, 2012).

$$F(u, v) = \sqrt{\frac{2}{N}} \sqrt{\frac{2}{M}} \sum_{N=0}^{N-1} \sum_{M=0}^{M-1} \Delta(i) \Delta(j) \cos \left[\frac{\pi u}{2N} (2i + 1) \right] \cos \left[\frac{\pi v}{2M} (2j + 1) \right] f(i, j) \quad (1)$$

$$\Delta(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{se } k = 0 \\ 1 & \text{outros} \end{cases} \quad (2)$$

Uma abordagem tradicional utilizada para simplificar a DCT-2D é utilizar a propriedade da separabilidade (GHANBARI, 2003). A separabilidade consiste em aplicar duas DCT-1D sucessivamente para obter a DCT-2D. Por consequência a DCT-2D é composta de dois passos consecutivos interligados por uma etapa de transposição.

Na etapa das transformadas diretas, devem-se calcular todos os tamanhos de DCTs para um quadro, e posteriormente escolher os blocos que obtiveram os melhores resultados, com isso, conforme vai diminuindo o tamanho do bloco manipulado pela DCT, temos que processar quatro vezes mais blocos, ou seja, considerando um quadro de um vídeo com resolução HD1080p, temos que processar 2010 blocos para transformadas de 32x32, 8040 blocos para 16x16, 32400 blocos para 8x8 e 129600 blocos para 4x4.

Com o objetivo de procurar reduzir a quantidade de blocos que são processados no decorrer da codificação, procurou-se realizar testes com várias configurações admissíveis pelo código de referência do HEVC, o *HEVC Model* (HM). Os testes realizados consistiram na variação da restrição do uso dos diferentes tamanhos de blocos utilizados pelas transformadas, onde foram considerados aceitáveis os resultados de redução da quantidade de blocos que não deteriorassem significativamente a compressão e a qualidade do vídeo.

Para cada perfil de teste foi efetuada a contagem da quantidade de blocos de cada tamanho de bloco processados durante o processo de codificação. Após o processo de codificação do vídeo teste, foi analisada a métrica *BD-Rate*, esta que relaciona a taxa de compressão em relação ao nível de qualidade do vídeo codificado. Ao todo 9 perfis de restrições foram utilizados, sendo que cada um destes perfis utiliza algum nível de restrição no tamanho de bloco e além dos 9 perfis mais um perfil foi utilizado que é padrão do codificador. Para simulação foi utilizado o vídeo Kimono com resolução HD1080p, contendo 240 frames e uma taxa de 24 quadros por segundo.

3. RESULTADOS E DISCUSSÃO

Na Tabela 1 podemos ver os resultados de *BD-rate* para cada teste realizado. Sendo o perfil A, o padrão do codificador, o mesmo não possui *BD-Rate*, pois a métrica é calculada a partir da variação de PSNR e Bitrate entre duas simulações distintas. Valores baixos de *BD-Rate* indicam que o perfil um impacto pequeno na quantidade de bits necessários para codificar o vídeo. De todos os perfis analisados, o que apresentou o menor acréscimo de taxa de bits para uma dada qualidade foi o perfil E, sendo seu *BD-Rate* de 0,4338%. Este perfil possui a vantagem de não processar transformadas diretas com blocos de 4x4 amostras de luminância, o que reduz significativamente a quantidade de blocos a serem processados.

Tabela 1. Resultados de *BD-Rate*

Perfil	A	B	C	D	E	F	G	H	I	J
	32x32 a 4x4	16x16 a 4x4	8x8 a 4x4	4x4	32x32 a 8x8	16x16 a 8x8	8x8	32x32 a 16x16	16x16	32x32
BD-Rate (%)	-	4,177	15,381	34,239	0,4338	4,582	15,952	6,074	11,432	22,219

Um ponto importante a ser mencionado é parâmetro de quantização (QP). O QP está relacionado com o módulo de quantização, sendo ele responsável pela maior taxa de compressão do vídeo. Ele basicamente corta os coeficientes colocados eles para zero, assim quanto maior o QP mais ele leva a zero os coeficientes.

Na Tabela 2 é apresentada a porcentagem de redução de blocos entre os perfis A e E, considerando os quatro valores de QP tipicamente utilizados no HEVC. Os coeficientes positivos representam que o número de blocos processados do respectivo tamanho foi reduzido.

Analisando a Tabela 2, constata-se que com QP 37 foi obtido uma redução de 73% nas transformadas de tamanho 4x4 e de 91% nas transformadas inversas. Continuando a análise nas transformadas de tamanho de 8x8, também tivemos uma redução na quantidade de 11% nas diretas e 16% nas inversas. Nas transformadas de 16x16 ocorreu um aumento de 1,2% na direta e 4,1% de redução nas inversas.

Tabela 2. Análise percentual da redução da quantidade de blocos processados

Tamanho	Redução (%)											
	4x4			8x8			16x16			32x32		
	Coder		Decoder	Coder		Decoder	Coder		Decoder	Coder		Decoder
QP	Direta	Inversa	Inversa	Direta	Inversa	Inversa	Direta	Inversa	Inversa	Direta	Inversa	Inversa
22	75,17	87,53	72,85	14,07	14,37	-12,68	2,91	4,53	-5,65	0,06	-0,12	-2,33
27	73,95	90,11	72,97	12,9	16,07	-7,51	1,62	4,74	-4,61	-0,01	-0,12	-0,13
32	73,05	91,51	72,27	12,05	16,63	-7,71	0,04	4,72	-4,19	-0,16	-0,15	0,82
37	73,06	91,46	64,02	11,96	16,22	-2,05	-1,27	4,16	-1,95	0,09	0,15	0,29

Com o QP 22, foi obtida uma redução de 75% nas transformadas diretas, e uma redução de 87% nas transformadas inversas no tamanho de bloco 4x4. Para blocos 8x8, uma redução de aproximadamente 14% nas diretas e nas inversas. Nas transformadas de 16x16 houve uma redução de 2,9% nas diretas e 4,5% nas inversas.

Considerando os demais QP, os valores de redução deles ficaram entre as faixas do QP de 22 e 37 ou ocorreu uma variação desprezível.

É possível observar que a restrição para uso de blocos 8x8 a 32x32 permite o processamento de blocos 4x4, entretanto em menor quantidade, quando comparado ao perfil padrão do HEVC. A explicação para este fato está no espaço de cores YCbCr utilizando e a subamostragem 4:2:0 (RICHARDSON, 2003), onde para cada 4 amostras de luminância (Y), há uma de crominância azul (Cb) e uma de crominância vermelha (Cr). Desta forma, conclui-se que a restrição dos tamanhos de blocos do HEVC estão relacionados aos blocos de amostras de luminância.

4. CONCLUSÕES

Neste trabalho foi apresentado um estudo inicial para redução da quantidade de informações processadas pelo módulo de transformação do padrão HEVC. Como pode ser observado, o perfil E (permite processar blocos de 8x8 a 32x32 amostras de luminância), apresentou-se promissor por possuir desprezível variação de qualidade e compressão em relação as configurações padrão do codificador.

A partir de uma análise mais detalhada na quantidade de blocos que foram reduzidos, foi possível constatar que as reduções são significativas, principalmente por apresentar maior redução nas transformadas de tamanho 4x4 e contribuindo para uma solução factível do codificador para dispositivos embarcados e/ou de baixo custo.

Como trabalhos futuros, pretende-se continuar está análise para os outros vídeos HD1080p que compõem as condições comuns de testes utilizadas na área de codificação de vídeo. Além disso, expandir a análise para resolução UHD 4K (3840x2160 pixels), de forma a verificar a possibilidade do não processamento de todos os possíveis tamanhos de blocos utilizados no padrão HEVC para esta resolução.

5. REFERÊNCIAS BIBLIOGRÁFICAS

JCT-VC Editors, **Recommendation ITU-T H.265 - High Efficiency Video Coding** (ITU-T Rec.H.265), April 2013. online. Disponível em: <http://handle.itu.int/11.1002/1000/11885>

International Telecommunication Union (ITU). **ITU-T Recommendation H.264/AVC :advanced video coding for generic audiovisual services**. janeiro 2012. online. Disponível em: <http://handle.itu.int/11.1002/1000/11466>

Bossen, F.; Bross, B.; Sühring, K., and D. Flynn, "HEVC complexity and implementation analysis," **IEEE Trans. Circuits Syst. Video Technol.**, vol. 22, no. 12, pp. 1684–1695, Dec. 2012.

Sullivan, G. J.; Ohm, J.-R.; Han, W.-J.; Wiegand, T. . "Overview of the High Efficiency Video Coding (HEVC) standard," **IEEE Trans. Circuits Syst. Video Technol.**, vol. 22, no. 12, pp. 1648–1667, Dec. 2012

GHANBARI, M. **Standard Codecs: Image Compression to Advanced Video Coding. United Kingdom**. The Institution of Electrical Engineers, 2003.

RICHARDSON, I. E-book. **The H.264 Advanced Video Compression Standard**. 2a.Ed. John Wiley & Sons, 2010.