

## ANÁLISE DA TAXA DE UTILIZAÇÃO DE TRANSFORMADAS EM QUADROS INTRA CODIFICADOS SEGUNDO O PADRÃO VVC

BIANCA SILVEIRA<sup>1</sup>; CLÁUDIO DINIZ<sup>2</sup>; DANIEL PALOMINO<sup>3</sup>; GUILHERME  
CORRÊA<sup>4</sup>;

<sup>1</sup>Universidade Federal de Pelotas – [bscsilveira@inf.ufpel.edu.br](mailto:bscsilveira@inf.ufpel.edu.br)

<sup>2</sup>Universidade Federal do Rio Grande do Sul – [cmdiniz@inf.ufrgs.br](mailto:cmdiniz@inf.ufrgs.br)

<sup>3</sup>Universidade Federal de Pelotas - [dpalomino@inf.ufpel.edu.br](mailto:dpalomino@inf.ufpel.edu.br)

<sup>4</sup>Universidade Federal de Pelotas – [gcorrea@inf.ufpel.edu.br](mailto:gcorrea@inf.ufpel.edu.br)

### 1. INTRODUÇÃO

A grande demanda de utilização de vídeos digitais vem crescendo numa velocidade surpreendente. Estima-se que até o ano de 2022 o tráfego de IP da internet chegue a 82% somente de transmissão de vídeo segundo GMDT FORECAST (2019). Assim, o processo de compressão de vídeo se torna de extrema importância para que esta transmissão seja realizada de forma eficiente.

O *Versatile Video Coding* (VVC) é um dos mais novos padrões de compressão de vídeo desenvolvido. Este padrão, que foi finalizado em julho de 2020, apresenta várias novas ferramentas de codificação que permitem até 40% de ganho de codificação em comparação com o seu antecessor *High Efficient Video Coding* (HEVC) SIDATY (2019). Certamente para que se possa obter essas vantagens de codificação, este padrão apresenta uma complexidade computacional das operações muito elevada, necessitando técnicas de otimizações algorítmicas e implementações de hardware de baixo consumo.

O padrão VVC apresenta inovações nos variados módulos do codificador, sendo uma delas o tipo de particionamento dos blocos. Para a realização das predições, os quadros podem ser divididos em tamanhos de blocos que variam de 4x4 até 128x128, podendo ser também de tamanhos retangulares como  $N \times N/2$ ,  $N/2 \times N$ ,  $N \times N/4$  e  $N/4 \times N$ . Outra novidade do padrão é a introdução da ferramenta *Multiple Transform Selection* (MTS) nas etapas de predição Intra e Inter do padrão. A MTS é composta por três tipos de transformadas: transformadas discretas do cosseno tipo II e tipo VIII (DCT-II e DCT-VIII) e a transformada discreta do seno tipo VII (DST-VII) com tamanhos de bloco que podem chegar a 64 x 64 para DCT-II e 32 x 32 para DCT-VIII/DST-VII. Além da MTS, o módulo das transformadas do VVC, pode realizar transformações nos blocos retangulares oriundos das predições e aplicar uma segunda transformada antes do processo de quantização.

Para usufruir destes novos recursos é necessário a utilização de técnicas que permitam acelerar o processo de codificação e assim, reduzir a complexidade das operações. O suporte às múltiplas transformadas encontradas na ferramenta MTS possui várias consequências relacionadas aos recursos de memória e lógica de alocação, portanto fornecer um projeto de alto desempenho para as restrições de hardware se torna uma questão crucial segundo Kammoun (2019).

Neste trabalho é realizada uma análise de usabilidade da MTS no software de referência VVC *Test Model* (VTM 10.0) VTM (2020) do padrão. O objetivo é extrair informações de quantas vezes estas transformadas são testadas para os blocos escolhidos na predição intra do codificador, permitindo determinar quais transformadas são prioritárias para uma futura implementação em hardware do

módulo MTS. Para se obter estas informações, será realizada uma análise dinâmica *profiling* do software de referência VTM 10.0.

## 2. METODOLOGIA

A *Multiple Transforms Selection* (MTS) utiliza mais de um tipo de transformadas para realizar a codificação residual dos blocos inter e intra. Ao padrão VVC foram acrescentadas mais duas transformadas, a transformada discreta do cosseno tipo 8 e a transformada discreta do seno tipo 7.

Para cada bloco, uma transformada na horizontal em uma dimensão (1D) é aplicada, obtendo assim um bloco  $N \times N$  intermediário. Novamente uma transformada 1D na vertical é aplicada sobre este bloco intermediário, resultando em um bloco final transformado. O que caracteriza a MTS é que os tipos de transformadas utilizadas nestas 2 dimensões podem ser distintas e estas são escolhidas pelo codificador.

Para realizar este controle da MTS, algumas *flags* de habilitação são especificadas em um conjunto de parâmetros de sequência de vídeo (*Sequence Parameter Set* - SPS) encontrados no software de referência VTM 10.0. Esta flag da MTS é habilitada como mostra o esquema da Tabela 1.

Tabela 1: Tabela de Mapeamento das Transformadas e Sinalização.

MTS_flag	Hor_flag	Ver_flag	Intra/Inter	
			Horizontal	Vertical
0			DCT-II	
1	0	0	DST-VII	DST-VII
	0	1	DCT-VIII	DST-VII
	1	0	DST-VII	DCT-VIII
	1	1	DCT-VIII	DCT-VIII

Tabela 2: Índice da MTS

MTS IDX	Horizontal	V ertical
MTS 0	DCT-II	DCT-II
MTS 1	SKIP	
MTS 2	DST-VII	DST-VII
MTS 3	DCT-VIII	DST-VII
MTS 4	DST-VII	DCT-VIII
MTS 5	DCT-VIII	DCT-VIII

Devido à adição de mais tipos de transformadas no padrão VVC, a etapa de transformadas demanda de um esforço computacional maior. Para validar estas afirmações, neste trabalho foi realizada uma análise da taxa de usabilidade das transformadas com a ferramenta MTS apenas para a etapa de predição Intra.

Para a realização deste estudo, foram testadas cinco sequências de vídeos, as quais podem ser encontradas no conjunto de testes *Common Test Conditions* (CTC) BOYCE (2018). A análise foi realizada com a configuração do SPS *All-intra* e com a *flag TemporalSubsampleRatio* igual a oito. Esta *flag* nos indica que um quadro é codificado a cada oito quadros. Foram utilizados 5 frames de cada sequência de vídeo.

No software de referência é possível encontrar um parâmetro chamado MTS-IDX (MTS *Index*), o qual determina qual o conjunto de transformadas que será utilizado para as operações. Na Tabela 2 pode se ver a configuração deste parâmetro. Se o MTS-IDX for igual a 0, o conjunto de transformadas escolhidas será a DCT-II aplicada na horizontal e vertical. Caso seja igual a 1, será aplicada a transformada SKIP (nenhuma transformada) e assim sucessivamente para os outros índices.

Para a extração das informações desejadas foi necessária uma manipulação do software de referência VTM. Primeiro foi localizado onde é realizada a escolha da melhor *Coding Structure* candidata, onde serão testadas as transformadas

para os diferentes tamanhos de blocos. Após, foram introduzidos contadores para ver o número de vezes que este índice é chamado.

### 3. RESULTADOS E DISCUSSÃO

Como dito anteriormente, a taxa de usabilidade das transformadas do padrão VVC foi dada por quantas vezes cada conjunto de transformadas foi testado para a melhor candidata de *Coding Structure*. Uma normalização foi aplicada aos valores obtidos para realizar uma comparação justa entre os diferentes tamanhos de blocos. Para esta técnica foi utilizado como referência o número de amostras presentes em um bloco 64x64 (4096 amostras).

No gráfico da Figura 1 é apresentada a taxa de quantas vezes foi testado um grupo de transformadas da ferramenta MTS. Os valores apresentados correspondem à taxa média das cinco sequências de vídeo analisadas.

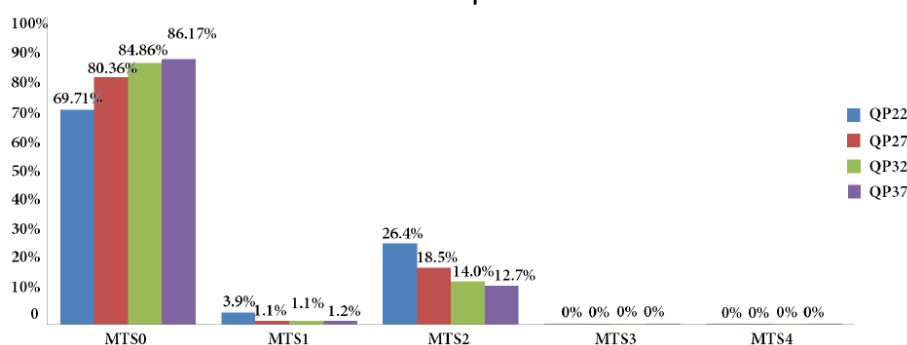


Figura 1. Total Médio das 5 sequências testadas

Uma análise mais específica da utilização da MTS foi realizada nos tamanhos de blocos quadrados como é visto nos gráficos das Figura 2.

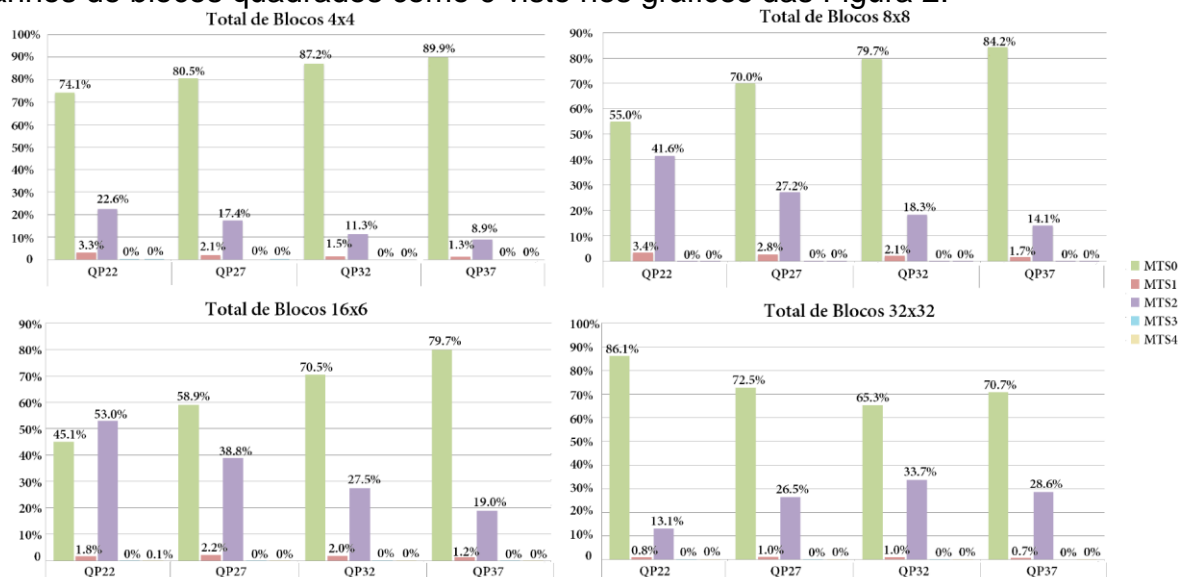


Figura 2. Total de blocos 4x4, 8x8, 16x16 e 32x32 testados para a melhor CS

Os resultados preliminares utilizando diferentes sequências de vídeo (com variações nas suas informações espaciais e temporais), mostram que a MTS0 (DCT-II na horizontal e vertical) é a mais utilizada para quase todos tamanhos de blocos e QP's com exceção do bloco 16x16 para QP igual a 22 que a MTS2 se apresentou mais significativa. A MTS0 apresentou uma taxa de 69,7%, 80,4%, 84,9% e 86,2% para os QP's 22, 27, 32 e 37, respectivamente. Para os tamanhos de blocos 4x4, 8x8 e 16x16, quanto maior é o QP, maior é a taxa de utilização da

MTS0. Já para estes mesmos tamanhos de blocos, a taxa de utilização da MTS2 aumenta enquanto os QP's diminuem.

Estes resultados preliminares apresentam apenas uma parte da taxa de utilização das transformadas para o codificador como um todo. Algumas explorações quanto ao uso das transformadas ainda são possíveis, como expandir esta análise para tamanhos de blocos retangulares, a análise de qual transformada está sendo escolhida como a melhor, investigando o custo de cada tamanho de bloco para cada posição de um quadro, etc.

Por ser um padrão muito recente, é muito difícil encontrar trabalhos relacionados ao seu uso e ferramentas. Os resultados sugerem que a ideia de desenvolver e implementar arquiteturas eficientes de hardware usando mais de um tipo de transformada permitiria otimizar os processos de codificação.

#### 4. CONCLUSÕES

Este trabalho apresentou uma análise da taxa de utilização das transformadas da ferramenta MTS na predição intra do novo padrão de codificação de vídeo *Versatile Video Coding*. A introdução deste novo recurso resulta em um maior número de operações pelo codificador e conseqüentemente em um maior esforço computacional, portanto uma investigação da taxa de usabilidade desta ferramenta foi realizada no software de referência VTM 10.0. Os resultados apresentaram que transformada mais utilizada é a DCT-II, utilizada na vertical e horizontal, seguida pela DST-VII (horizontal e vertical). A MTS0 apresentou uma taxa de utilização de 69,7%, 80,4%, 84,9% e 86,2% para os QP's 22, 27, 32 e 37, respectivamente. Considerando os trabalhos existentes na literatura, se evidencia uma escassez de projetos relacionados às transformadas do padrão VVC. Tendo em vista estes resultados e a atender às necessidades destas aplicações se torna importante o estudo de técnicas para otimização e implementações em hardware.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

SIDATY, N.; HAMIDOUCE, W.; D'EFORGES, O.; PHILIPPE, P.; FOURNIER, J. Compression performance of the versatile video coding: Hd and Uhd visual quality monitoring. **Picture Coding Symposium (PCS)**, IEEE, pp. 1-5, 2019.

KAMMOUN, A., HAMIDOUCE, W., PHILIPP, P., BELGHITH, F., MASSMOUDI, N., & NEZAN, J. F. Hardware acceleration of approximate transform module for the versatile video coding standard. **European Signal Processing Conference (EUSIPCO)** pp. 1-5, IEEE, 2019.

BOYCE, J.; SUEHRING, K.; Li, X.; **Jvet-j1010: Jvet common test conditions and software reference configurations**, JVET, San Diego, CA, USA, Tech. Rep. JVET-J1010, 2018.

GMDT FORECAST. **Cisco visual networking index: global mobile data traffic forecast update**, 2017-2022, vol. 2017, pp. 2022, 2019.

FRAUNHOFER, **Vtm reference software**, vvc test model (vtm), 2020.