

Eliminação de Candidatos Baseada em Taxa para Algoritmos de Estimação de Movimento na Codificação de Vídeo

Gabriel B. Sant’Anna, Luiz Henrique Cancellier, Ismael Seidel,
Mateus Grellert, José Luís Güntzel

Embedded Computing Lab. (ECL) – Dep. de Informática e Estatística (INE)
Universidade Federal de Santa Catarina (UFSC) – Florianópolis – Brasil

{baiocchi.gabriel, luizhenriquecancellier, ismaelseidel}@gmail.com,
{mateus.grellert, j.guntzel}@ufsc.br

Abstract. *This work proposes and evaluates a search space reduction technique for Motion Estimation, which is considered as one of the most important steps in video encoding. Based on previous works’ indications that encoder decisions are biased towards candidates with a lower vector cost, the proposed algorithm uses a simple elimination criterion based on the estimated bitrate of candidate motion vectors. In one of the tested configurations of the VVC reference encoder implementation, the technique was able to reduce Motion Estimation search area by more than 80% at the cost of an average 0.74% decrease in coding efficiency with respect to the original search algorithm.*

Resumo. *Este trabalho propõe e avalia uma técnica dedicada à redução do espaço de busca da Estimação de Movimento, uma das etapas mais importantes na codificação de vídeo. Partindo de evidências na literatura que indicariam uma correlação entre o custo estimado da representação de vetores de movimento e as decisões do codificador, o algoritmo proposto utiliza um critério de eliminação de candidatos considerando somente a taxa de bits dos seus respectivos vetores. A técnica foi avaliada no software de referência do padrão VVC e demonstrou, em uma das configurações do codificador, reduções de mais de 80% na área de busca da Estimação de Movimento, com perdas médias na eficiência de codificação de apenas 0,74% em relação ao algoritmo original.*

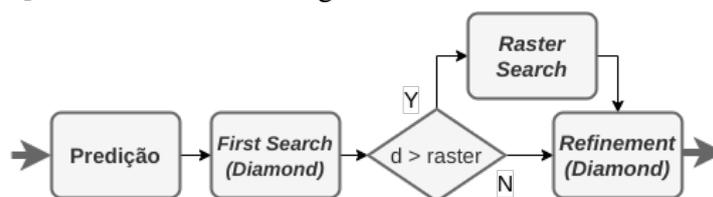
1. Introdução

Vídeos representam a maior parte de todo o tráfego da internet: em 2017 compunham aproximadamente 75% dos dados em rede, com indicações de que essa parcela chegaria a 82% até 2022 [Cisco 2019]. Tal fato advém da grande demanda por aplicações multimídia e vídeos em alta resolução, o que aumentou ainda mais – e em escala mundial – durante a pandemia da COVID-19 [Watson 2020]. Atualmente, é essa a tecnologia que permite a continuidade das atividades de trabalho e de ensino à distância durante o distanciamento social. Evidenciada a importância dos vídeos na comunicação e a intensificação do volume de dados dos mesmos, destaca-se a necessidade de técnicas de codificação cada vez mais eficientes – com maiores taxas de compressão para a mesma qualidade visual.

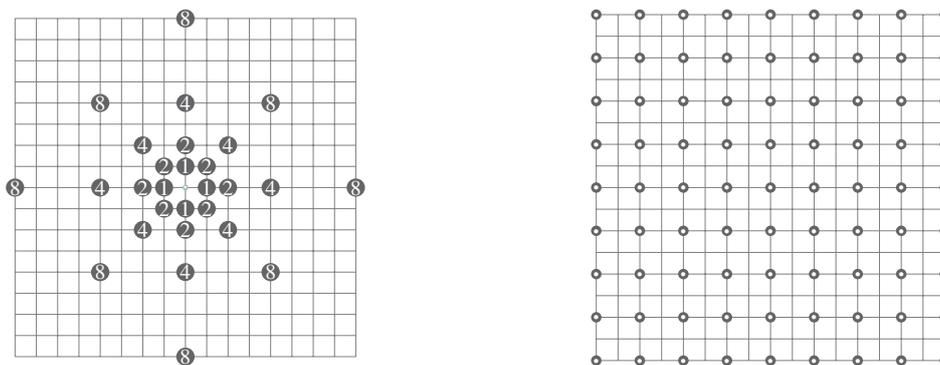
Para tratar vídeos em resoluções crescentes e atingir os níveis de compressão necessários para atender requisitos de banda diversos, padrões de codificação de vídeo estão em constante desenvolvimento. Entretanto, melhorias na eficiência de codificação são

geralmente acompanhadas por um maior tempo de execução do processo: o *Versatile Video Coding* (VVC) é o mais recente padrão da ITU-T e apresenta uma redução de *bitrate* média de 44,4% a custo de um tempo de codificação cerca de 10 vezes maior em relação ao padrão anterior, o *High Efficiency Video Coding* (HEVC) [Siqueira et al. 2020]. Este fenômeno pode chegar a impedir a utilização dos novos padrões em aplicações como o *streaming* em tempo real e a codificação de vídeo em dispositivos móveis, motivando o desenvolvimento de algoritmos rápidos.

A Estimação de Movimento – *Motion Estimation* (ME) – é considerada uma das partes mais importantes na compressão de vídeo e, com um algoritmo de busca exaustivo (*Full Search*), pode tomar até 80% de todo o tempo da codificação [Zhang et al. 2018]. Alternativamente, algoritmos rápidos abdicam de um resultado ótimo na busca para reduzir o tempo de execução do processo, o que costuma ser feito através de reduções na região de busca considerada durante a ME [Amirpour et al. 2019]. O algoritmo *Test Zone Search* (TZS) é um destes, destacando-se por ter sido adotado como o método padrão para a *Integer Motion Estimation* (IME) nas implementações de referência do HEVC [Rosewarne et al. 2019] e do VVC [Chen et al. 2019] devido à sua capacidade de evitar mínimos locais através da combinação de um variante da *Diamond Search* [Shan Zhu and Kai-Kuang Ma 2000] com a busca semi-exaustiva da *Raster Search* [Doan et al. 2017], conforme ilustra a Figura 1.



(a) Fluxograma de alto nível do algoritmo da TZS.



(b) Padrões de busca utilizados na TZS: *Diamond* (à esquerda) e *Raster* (à direita).

Figura 1. Esquemática do algoritmo TZS.

Embora a TZS melhore o desempenho da IME quando comparada com uma busca exaustiva, o número de candidatos avaliados ainda é considerado alto demais para aplicações de compressão de vídeo de tempo real [Zhang et al. 2018] [Fan et al. 2017]. Assim, com o objetivo de reduzir ainda mais o tempo de execução da IME, [Goncalves et al. 2018] sugerem utilizar um padrão de busca octogonal cuja forma foi definida levando em conta a distribuição média dos Vetores de Movimento – *Motion Vectors* (MVs) – selecionados pelo codificador. Infelizmente, a busca octogonal foi definida para atuar em um passo específico da TZS e não é facilmente combinada com outros

algoritmos rápidos. Enquanto isso, um outro trabalho correlato utiliza uma abordagem diferente para reduzir o tempo de execução da IME: ao se aproveitarem do fato de que, no método de *Successive Elimination* [Li and Salari 1995], a chance de eliminar um candidato aumenta quando o seu custo de MV é maior, [Trudeau et al. 2018] produziram uma variação do algoritmo que ordena os candidatos pela estimativa de taxa de bits dos seus vetores. O algoritmo resultante garante um resultado ótimo, mas no pior caso ainda considera todos os candidatos da região de busca.

Observando os trabalhos de [Goncalves et al. 2018] e [Trudeau et al. 2018], percebe-se que a distribuição de MVs selecionados, tal como a superfície da estimativa de taxa de MV, são fatores que mereceriam ser analisados ao tentar reduzir o espaço de busca da ME. Seguindo essas indicações, este trabalho traz as seguintes contribuições:

1. A identificação, na literatura, de padrões de busca para a IME que são marcadamente influenciados pela taxa de MVs, relação essa que nenhum trabalho anterior parece explorar de maneira explícita [Zhang et al. 2018].
2. A proposta de uma estratégia que utiliza o componente de taxa da função de custo como critério local para eliminar candidatos e reduzir a área de busca da IME.
3. Um estudo de caso aplicando a técnica proposta sobre a implementação da TZS do *VVC Test Model* (VTM) e uma comparação com o padrão de busca octogonal de [Goncalves et al. 2018].

Destaca-se também que este trabalho de iniciação científica originou um artigo aceito na *International Conference on Acoustics, Speech, and Signal Processing* (ICASSP) 2021.

O texto que segue está organizado da seguinte forma: a Seção 2 apresenta os conceitos da ME que serão relevantes para explicar o algoritmo na Seção 3. Na Seção 4, os experimentos realizados são detalhados e seus resultados discutidos. Por fim, a Seção 5 elenca algumas conclusões e sugere possíveis trabalhos futuros.

2. Fundamentação

De maneira geral, a compressão de vídeos é feita a partir da exploração de três tipos de redundâncias: as entrópicas, relacionadas à representação binária da informação; as espaciais, da semelhança entre regiões de um mesmo quadro; e as temporais, das similaridades entre diferentes quadros de um vídeo. O processo que explora as redundâncias temporais em vídeos é a predição inter-quadros, dentro do qual a ME é responsável por representar a informação de um bloco (região de um quadro) através da sua diferença com um segundo bloco – chamado bloco de referência, que se localiza em um quadro diferente do original – juntamente de um MV, que nada mais é do que a posição relativa entre eles. A Figura 2 ilustra a relação entre bloco original, bloco candidato e MV.

O processo é dividido em três etapas: Predição, ME Inteira (IME) e ME Fracionária (FME) [Zhang et al. 2018]. Primeiramente, a etapa de Predição utiliza a informação de movimento de blocos próximos do original para derivar um Vetor Preditor – *Predicted Motion Vector* (PMV) – que define o quadro candidato e o ponto de partida da busca que se segue. Em seguida, a IME busca, dentro de uma região centrada no PMV, pelo candidato que minimiza uma função de custo responsável por ponderar o resíduo resultante da diferença dos blocos e o custo da representação do MV entre eles. Por fim, a FME executa um processo de refinamento do resultado da busca inteira.

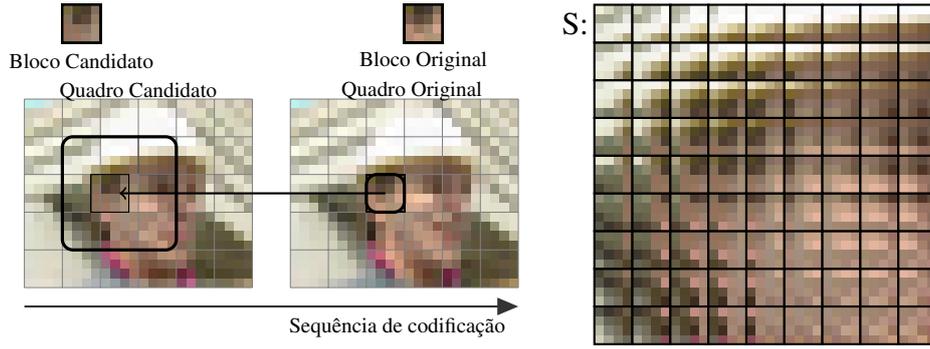


Figura 2. Ilustração da IME relacionando um bloco original com uma referência selecionada dentre os candidatos no conjunto de busca S , bem como o MV entre eles. Adaptado de [Cancellier et al. 2017].

A Equação (1) mostra a função de custo minimizada durante a IME [Sullivan and Wiegand 1998], onde r estima a taxa de bits necessária para representar a diferença entre o MV candidato (\vec{mv}) e o preditor ($p\vec{mv}$), λ é um multiplicador de Lagrange que pondera a taxa e d calcula a distorção entre o bloco original (O) – dado no início do processo da ME – e um bloco candidato que é apontado pelo MV de entrada.

$$j(\vec{mv}) = d(\mathbf{C}^{\vec{mv}}) + \lambda \cdot r(\vec{mv} - p\vec{mv}) \quad (1)$$

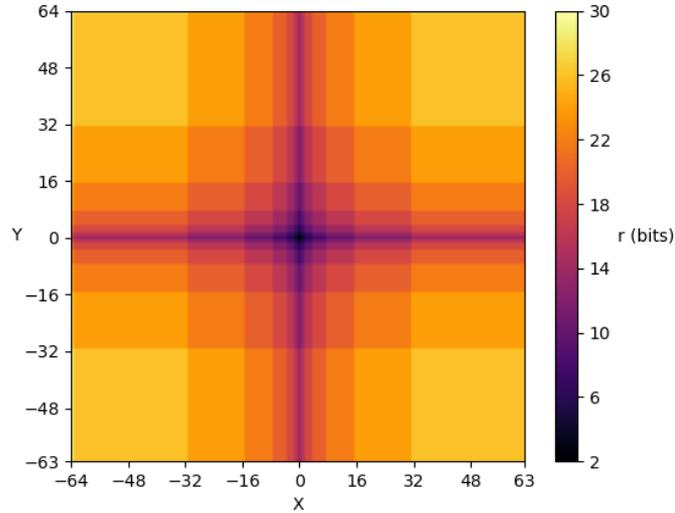


Figura 3. Superfície de taxa para coordenadas de MVs relativas a um preditor qualquer. Adaptado de [Sant'Anna et al. 2021].

A Soma das Diferenças Absolutas – *Sum of Absolute Differences* (SAD) – dada na Equação (2) é uma métrica de distorção comumente utilizada na codificação de vídeo. Apesar da sua expressão ser relativamente simples, a SAD é avaliada sobre cada candidato considerado durante a busca e requer a leitura desses blocos da memória, o que a torna a operação mais custosa da IME e inviabiliza buscas exaustivas.

$$d(\mathbf{C}) = \sum_{i=1}^m \sum_{j=1}^n |C_{i,j} - O_{i,j}| \quad (2)$$

Enquanto isso, a função que estima a taxa de bits necessária para representar um MV é implementada no VTM pela Equação (3), onde $g(v)$ denota o comprimento do

código Exp-Golomb para um valor inteiro e cada componente do vetor é dado relativo ao PMV. Destaca-se o potencial de otimização da computação de $r(\vec{mvd})$: diferentemente da distorção, sua expressão não depende do conteúdo do quadro sendo codificado e $g(v)$ pode ser pré-calculada pois a janela de busca limita os possíveis valores de mvd_x e mvd_y [Coban and Mersereau 1998]. A Figura 3 mostra a superfície determinada pela função de taxa em uma janela de 128×128 pixels centrada no PMV.

$$r(\vec{mvd}) = g(\vec{mvd}_x) + g(\vec{mvd}_y) \quad (3)$$

3. Eliminação de Candidatos Baseada em Taxa

Apesar da TZS reduzir significativamente o tempo de execução da IME em comparação com uma busca exaustiva, trabalhos na literatura já demonstraram que é possível avaliar menos candidatos ainda e manter níveis aceitáveis de eficiência de codificação. São especialmente relevantes neste contexto os trabalhos sobre a ME no HEVC que mostram que os MVs selecionados são geralmente aqueles próximos do preditor: [Goncalves et al. 2017] indicam que 87% dos candidatos são escolhidos na etapa de predição da TZS e [Fan et al. 2017] mostram que até 94% dos melhores MVs podem ser encontrados em uma pequena região de busca centrada no PMV.

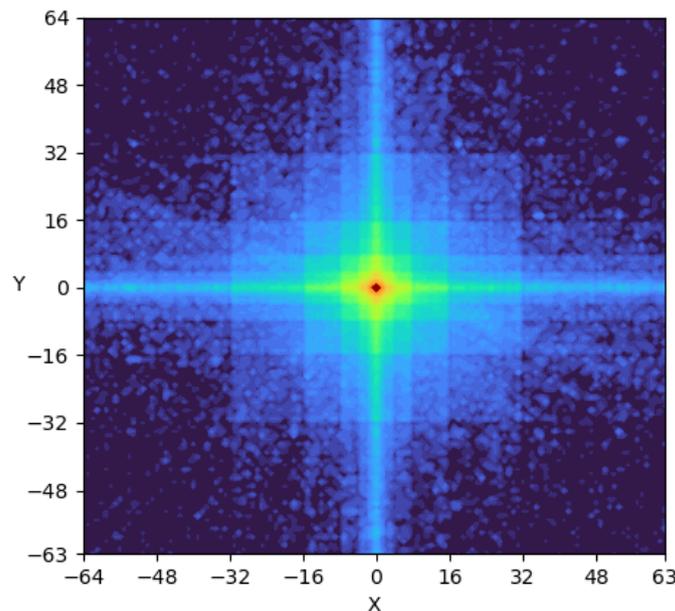


Figura 4. Distribuição média de MVs selecionados pelo VTM ao codificar os vídeos de teste (vide Seção 4), onde cores mais quentes indicam maior incidência de decisões naquele ponto. Adaptado de [Sant’Anna et al. 2021].

Para ilustrar esse ponto, a Figura 4 mostra a distribuição de MVs escolhidos pelo VTM. Ao parear, para cada posição de vetor, as taxas de bits da Figura 3 e os valores – em escala logarítmica – da Figura 4, encontra-se um coeficiente de correlação de Pearson de $-0,89$, indicando que o número de MVs selecionados diminui exponencialmente à medida que sua estimativa de taxa aumenta. Assim, conclui-se que a maioria das decisões estão dentro das regiões da área de busca onde a estimativa de taxa de bits de vetores é menor – coincidindo com as proximidades do PMV.

Partindo dessas observações, a técnica proposta consiste em reduzir o tempo de execução médio da função de custo ao longo da IME através da eliminação do cálculo da distorção para os MVs para os quais vale o critério da Equação (4), onde t é um valor inteiro que atua como limiar de taxa. Tendo em vista que computar o critério de eliminação tem um custo computacional desprezível em relação ao cálculo da distorção, pode-se considerar que a região de busca é efetivamente reduzida a uma área em forma de diamante – com diâmetro proporcional ao limiar t – centrada no preditor. A abordagem adotada é então justificada pela correlação entre a superfície de taxa e a distribuição de decisões do codificador, que possibilita eliminar a maior parte da janela de busca e ao mesmo tempo manter as regiões com maiores chances de conter os melhores candidatos.

$$r(\vec{m}v - p\vec{m}v) > t \quad (4)$$

Como possível desvantagem da técnica de eliminação de candidatos baseada em taxa, pode-se apontar a sua grande dependência na eficácia da etapa de predição, responsável por determinar o PMV.

4. Experimentos

Para avaliar a técnica descrita na seção anterior, o algoritmo da TZS do VTM 6.2 foi modificado¹ para aplicar a eliminação de candidatos baseada em taxa. A partir das *Common Test Conditions* (CTC) – as mesmas condições utilizadas para testar os padrões de codificação de vídeo da ITU-T – foi então definida uma série de experimentos para cada combinação de configuração em {Low Delay P (LDP), Random Access (RA)}, Quantization Parameter (QP) em {22, 27, 32, 37} e sequência de teste – cada um dos vídeos no subconjunto comum à CTC do HEVC [Bossen 2013] e do VVC [Bossen et al. 2019]. Em seguida, os experimentos foram executados em cada uma das versões do codificador: TZS original, TZS com eliminação de candidatos e, mais tarde, TZS com a busca octogonal de [Goncalves et al. 2018], totalizando assim mais de 400 testes.

Adotou-se como métrica de eficiência de codificação o *Bjontegaard Delta Bitrate* (BD-Rate) relativo ao algoritmo original, onde valores de BD-Rate positivos indicam que o vídeo passa a ocupar mais espaço para a mesma qualidade visual [Bjontegaard 2001]. Já a medida do efeito de um algoritmo sobre a área de busca da IME foi definida da seguinte forma: seja C a área de busca efetiva de um vídeo definida como na Equação (5), a redução percentual ΔC depende da área de busca da implementação original (C_{ori}) e da implementação modificada (C_{mod}), como indica a Equação (6).

$$C = \sum_{s \in S} totalCandidates(s) \times area(s) \quad (5)$$

$$\Delta C = \frac{C_{ori} - C_{mod}}{C_{ori}} \times 100 \quad (6)$$

Aqui, S é o conjunto de todos os possíveis tamanhos de bloco considerados pelo codificador, $totalCandidates(s)$ é o número de candidatos de dimensão s sobre os quais foi calculada a distorção durante a IME e $area(s)$ é a área daquele tamanho de bloco. Essa métrica foi escolhida por permitir a replicabilidade dos resultados em diferentes configurações de hardware, sistema operacional e compilador.

¹Implementação disponível em gitlab.com/baioc/vtm/tree/rate-elimination

Sendo $t = 4$ o menor valor para o qual a TZS testa outros candidatos além do PMV, esse foi o limiar de taxa utilizado para avaliar o algoritmo de eliminação de candidatos, cujos resultados são resumidos na Figura 5.

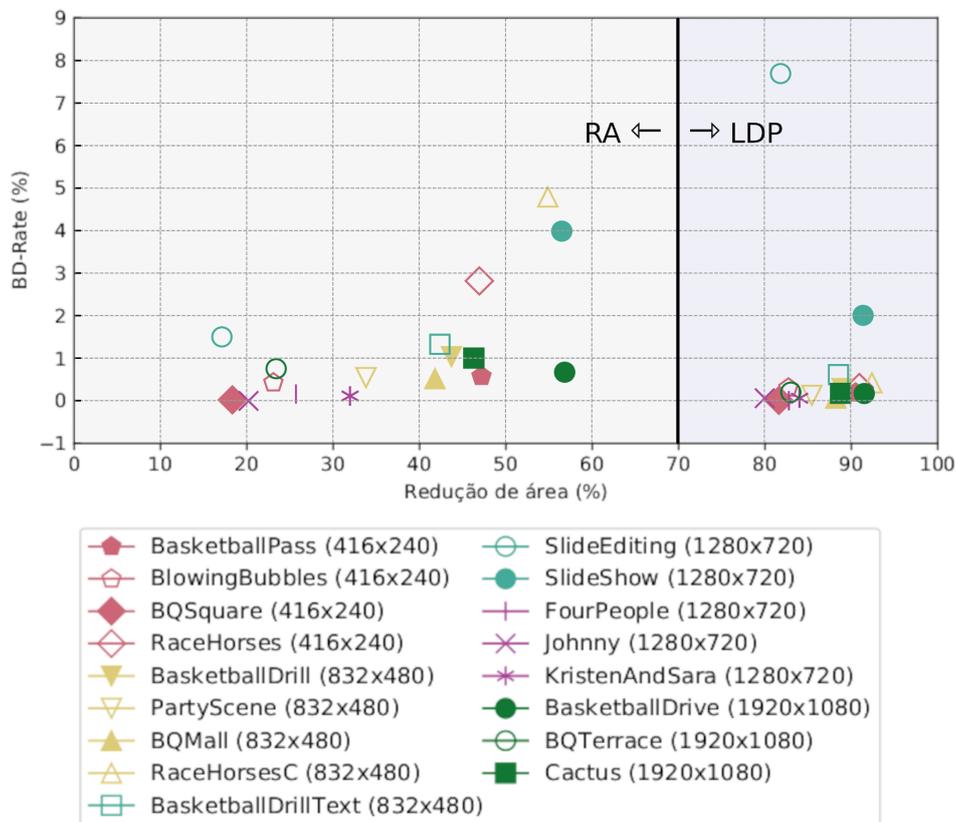


Figura 5. Resultados da eliminação de candidatos baseada em taxa aplicada, com um limiar fixo de $t = 4$, sobre a TZS. Inclui os resultados de ambas as configurações RA e LDP. Adaptado de [Sant’Anna et al. 2021].

Embora a maioria das sequências de teste tenham apresentado valores de BD-Rate inferiores a 1% (nível de perda considerado significativo em aplicações com restrições de banda conservadoras), os resultados mostram que $t = 4$ restringe excessivamente a busca para a configuração RA. Logo, sabendo que uma região de busca mais extensa deve resultar em menor BD-Rate, foram conduzidos experimentos adicionais relaxando o critério de eliminação para abaixar os níveis de BD-Rate até menos de 1%. Notavelmente, um limiar de $t = 20$ resulta em uma região de busca efetiva similar ao “octógono” desenvolvido por [Goncalves et al. 2018]. Assim, visando comparar a técnica de eliminação de candidatos com o trabalho de [Goncalves et al. 2018], o seu padrão de busca – originalmente implementado no HEVC – foi replicado no software do VTM e adaptado para diferentes dimensões de janela de busca.

A Tabela 1 mostra os resultados obtidos ao aumentar o limiar de taxa na configuração RA para os vídeos que anteriormente obtiveram BD-Rate acima de 1%. São mostrados também os valores relevantes obtidos dos experimentos com o padrão de busca octogonal. Como esperado, foi possível configurar o limiar de taxa em $t = 20$ para fazer com que a eliminação de candidatos alcançasse níveis de BD-Rate próximos daqueles exibidos pelo padrão octogonal.

Tabela 1. BD-Rate e redução da área de busca para diferentes valores de limiar de taxa na configuração RA e comparação com o padrão octogonal.

Sequência	$t = 10$		$t = 20$		Octagonal	
	BD-Rate (%)	ΔC (%)	BD-Rate (%)	ΔC (%)	BD-Rate (%)	ΔC (%)
<i>Cactus</i>	0,55	39,0	0,12	28,8	0,02	26,2
<i>BasketballDrill</i>	0,60	35,8	0,14	23,4	0,01	22,2
<i>BasketballDrillText</i>	0,67	34,3	0,09	22,9	-0,04	21,2
<i>SlideEditing</i>	1,23	10,0	0,63	6,8	0,03	6,1
<i>RaceHorses</i>	1,53	38,0	0,35	25,1	0,02	22,5
<i>SlideShow</i>	2,88	51,0	0,88	42,5	-0,05	36,4
<i>RaceHorsesC</i>	3,02	47,7	0,70	33,9	0,10	30,2

Por outro lado, o mesmo limiar de $t = 4$ traz resultados excelentes na configuração LDP: reduções na área de busca acima de 80%, com uma média de BD-Rate de 0,74%. A Tabela 2 divide os resultados nas classes definidas pela CTC e as compara com o padrão octogonal, também na configuração LDP. Observa-se que a comparação de redução da área de busca é dominada pelo algoritmo de eliminação de candidatos, com resultados médios de 86,69% enquanto a busca octogonal alcança apenas 16,4% no melhor caso. Ambos os algoritmos apresentam valores aceitáveis de BD-Rate, as únicas exceções sendo as sequências *SlideEditing* e *SlideShow* da classe F no caso do algoritmo proposto. Tais casos excepcionais são justificados pelo conteúdo não usual desses dois vídeos, capturas de tela de apresentações de slides, cujo tratamento adequado em codificadores modernos requer ferramentas especializadas.

Tabela 2. Médias, por classe da CTC, de BD-Rate e redução da área de busca em comparação com o padrão octogonal, na configuração LDP.

Classe	$t = 4$		Octagonal	
	BD-Rate (%)	ΔC (%)	BD-Rate (%)	ΔC (%)
B	0,18	87,8	0,02	13,9
C	0,22	88,8	0,00	15,2
D	0,20	86,5	0,04	10,7
E	0,04	82,3	-0,04	6,5
F	3,44	87,3	0,37	16,4

5. Conclusões

Neste trabalho, foi demonstrado como a estimativa de taxa de bits dos MVs pode influenciar os padrões de busca da IME. Reforçando o ponto anterior, foi desenvolvido um algoritmo que tira proveito dessa relação para eliminar partes da região de busca através de um critério simples e eficiente. As principais vantagens dessa técnica são sua versatilidade, já que pode ser combinada com outros algoritmos de busca existentes, e baixo custo computacional, visto que a superfície de taxa pode ser pré-computada em *lookup tables*.

Os experimentos realizados mostraram que mesmo um limiar de taxa fixo aplicado sobre a TZS pode reduzir a área de busca da IME em mais de 80%, em troca de um aumento de BD-Rate médio de apenas 0,74% na configuração LDP. Embora a relação entre redução de área e eficiência de codificação tenha sido menos interessante na configuração RA, foi possível exemplificar a parametrização da técnica desenvolvida

ao variar o limiar de taxa até alcançar uma região de busca similar àquela proposta por [Goncalves et al. 2018].

Trabalhos futuros poderiam propor modificações e melhorias dessa técnica, bem como estudar seu uso sobre outros algoritmos da ME. A identificação de casos de teste discrepantes, por exemplo, indica que um limiar adaptativo poderia ser usado para melhorar os resultados de BD-Rate ao acomodar as características de movimento de cada vídeo. Por fim, dado que um limiar de taxa fixo trouxe bons resultados na configuração LDP, implementações de codificadores com blocos de hardware dedicados (aceleradores) poderiam aplicar, estaticamente, um critério de eliminação (sobre uma busca exaustiva, por exemplo) para simplificar drasticamente o processo de busca da IME, reduzindo-a a uma pequena região em torno do preditor e possivelmente trazendo reduções significativas de área, custo e consumo energético da ME.

Prêmios e Publicações

Este trabalho foi desenvolvido pelo aluno de iniciação científica (primeiro autor) com a orientação e supervisão de outros membros do laboratório (demais autores). A pesquisa deu origem a um artigo de conferência, que foi aceito para apresentação oral e publicação nos anais da ICASSP 2021, evento internacional da *Institute of Electrical and Electronics Engineers* (IEEE) classificado como *qualis* A1. Além disso, o resumo e apresentação das atividades desenvolvidas obtiveram a melhor avaliação do departamento e se destacaram com colocação entre os dois melhores trabalhos da grande área (Ciências Exatas, da Terra e Engenharias) no 30º Seminário de Iniciação Científica e Tecnológica (SIC) da UFSC.

Referências

- Amirpour, H., Ghanbari, M., Pinheiro, A., and Pereira, M. (2019). Motion estimation with chessboard pattern prediction strategy. *Multimedia Tools and Applications*, 78(15):21785–21804.
- Bjøntegaard, G. (2001). Calculation of average PSNR differences between RD-curves. Document VCEG-M33, Video Coding Experts Group (VCEG), Austin, Texas, USA.
- Bossen, F. (2013). Common HM test conditions and software reference configurations. Document JCTVC-L1100, JCT-VC of ITU-T, Geneva.
- Bossen, F., Boyce, J., Li, X., Seregin, V., and Sühring, K. (2019). JVET common test conditions and software reference configurations for SDR video. Document JVET-N1010, JVET of ITU-T, Geneva.
- Cancellier, L., Seidel, I., and Güntzel, J. L. (2017). Algoritmo de eliminações sucessivas em níveis baseado na soma das diferenças transformadas absolutas. In *Anais do XXXVI Concurso de Trabalhos de Iniciação Científica da SBC*, Porto Alegre, RS, Brasil. SBC.
- Chen, J., Ye, Y., and Kim, S. (2019). Algorithm description for Versatile Video Coding and Test Model 6 (VTM 6). Document JVET-O2002, JVET of ITU-T, Gothenburg.
- Cisco (2019). Cisco Visual Networking Index: Forecast and Trends, 2017–2022 White Paper. <https://web.archive.org/web/20190820234616/https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html>. Accessed: 20 Aug. 2019.

- Coban, M. Z. and Mersereau, R. M. (1998). A fast exhaustive search algorithm for rate-constrained motion estimation. *IEEE Transactions on Image Processing*, 7(5):769–773.
- Doan, N., Kim, T. S., Rhee, C. E., and Lee, H.-J. (2017). A hardware-oriented concurrent TZ search algorithm for High-Efficiency Video Coding. *EURASIP Journal on Advances in Signal Processing*, pages 1–17.
- Fan, R., Zhang, Y., and Li, B. (2017). Motion Classification-Based Fast Motion Estimation for High-Efficiency Video Coding. *IEEE Transactions on Multimedia*, 19(5):893–907.
- Goncalves, P., Correa, G., Porto, M., Zatt, B., and Agostini, L. (2017). Multiple early-termination scheme for TZ search algorithm based on data mining and decision trees. In *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–6.
- Goncalves, P., Porto, M., Zatt, B., Agostini, L., and Correa, G. (2018). Octagonal-Axis Raster Pattern for Improved Test Zone Search Motion Estimation. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1763–1767.
- Li, W. and Salari, E. (1995). Successive elimination algorithm for motion estimation. *IEEE Transactions on Image Processing*, 4(1):105–107.
- Rosewarne, C., Sharman, K., Sjöberg, R., and Sullivan, G. J. (2019). High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description Update 12. Document JCTVC-AK1002, JCT-VC of ITU-T, Geneva.
- Sant’Anna, G. B., Cancellier, L. H., Seidel, I., Grellert, M., and Güntzel, J. L. (2021). Relying on a rate constraint to reduce motion estimation complexity. <https://arxiv.org/abs/2102.09656>. Preprint.
- Shan Zhu and Kai-Kuang Ma (2000). A new diamond search algorithm for fast block-matching motion estimation. *IEEE Transactions on Image Processing*, 9(2):287–290.
- Siqueira, I., Correa, G., and Grellert, M. (2020). Rate-Distortion and Complexity Comparison of HEVC and VVC Video Encoders. In *2020 IEEE 11th Latin American Symposium on Circuits Systems (LASCAS)*, pages 1–4.
- Sullivan, G. J. and Wiegand, T. (1998). Rate-distortion optimization for video compression. *IEEE Signal Processing Magazine*, 15(6):74–90.
- Trudeau, L., Coulombe, S., and Desrosiers, C. (2018). Cost-Based Search Ordering for Rate-Constrained Motion Estimation Applied to HEVC. *IEEE Transactions on Broadcasting*, 64(4):922–932.
- Watson, A. (2020). Consuming media at home due to the coronavirus worldwide. <https://www.statista.com/statistics/1106498/home-media-consumption-coronavirus-worldwide-by-country/>. Accessed: 16 Jun. 2020.
- Zhang, Y., Zhang, C., and Fan, R. (2018). Fast Motion Estimation in HEVC Inter Coding: An Overview of Recent Advances. In *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 49–56.