ICASSP2021 TORONTO Canada June 6-11, 2021

Relying on a rate constraint to reduce Motion Estimation complexity

#3609

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

Embedded Computing Laboratory (ECL) – Federal University of Santa Catarina (UFSC) – Florianópolis, Brazil

Introduction

- **Video data amounts to the majority of global internet traffic** [3], even more so during the COVID-19 pandemic
- In order to achieve ever higher compression rates, video encoders are becoming increasingly complex [7]
- When an exhaustive search is used, **Motion Estimation (ME) can be responsible for up to 80% of encoding time** [9]
- The widely adopted **Test Zone Search (TZS) algorithm is considered too slow for some applications** [4]

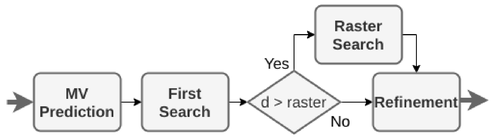

Figure 1: High-level flowchart of the TZS algorithm.

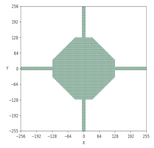

Figure 2: Octagonal-axis raster pattern, as replicated in VTM.

- Among alternatives to standard TZS, the Octagonal-axis [6] search pattern was designed using average Motion Vector (MV) encoder decisions
- We show how **this idea can be generalized using the motion vector bitrate to constrain the IME search area**

Rate-based Candidate Elimination

The cost function minimized during ME [8] is

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

- r estimates the bitrate of a candidate MV
- d is block distortion, the most costly computation in IME

Previous works have indicated that selected MVs:

- Can usually be found right after the TZS prediction step [5]
- Are mostly within a small area around the predictor [4]

Contribution

We define a bitrate threshold t and use it to constrain the IME search area, eliminating distortion calculation whenever

$$r(\vec{mv} - \vec{mvp}) > t$$

- **Simple** and **efficient** elimination criterion
- Skipped blocks need not be fetched from memory
- **Compatible** with existing Rate-constrained ME algorithms
- **Configurable** through the bitrate elimination threshold

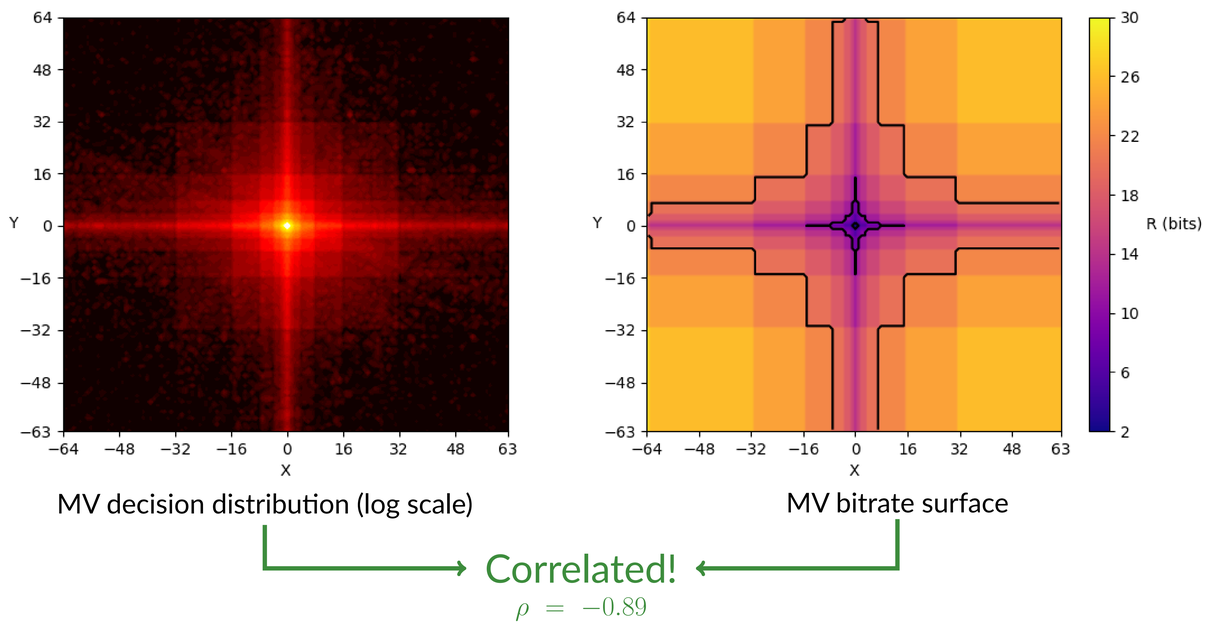

Figure 3: Correlation between MV decision distribution and MV bitrate surface.

Experimental Results

Our proposed candidate elimination technique was evaluated by combining:

- The VTM 6.2 encoder
- 17 test sequences from the Common Test Conditions [1] [2]
- QPs 22, 27, 32 and 37
- Encoder configurations LDP and RA

Encoding efficiency was measured in BD-Rate and ME complexity estimated through block area for which distortion was computed:

$$\Delta C = \frac{C_{ori} - C_{mod}}{C_{ori}} \times 100\% \qquad C = \sum_{s \in S} totalCandidates(s) \times area(s)$$

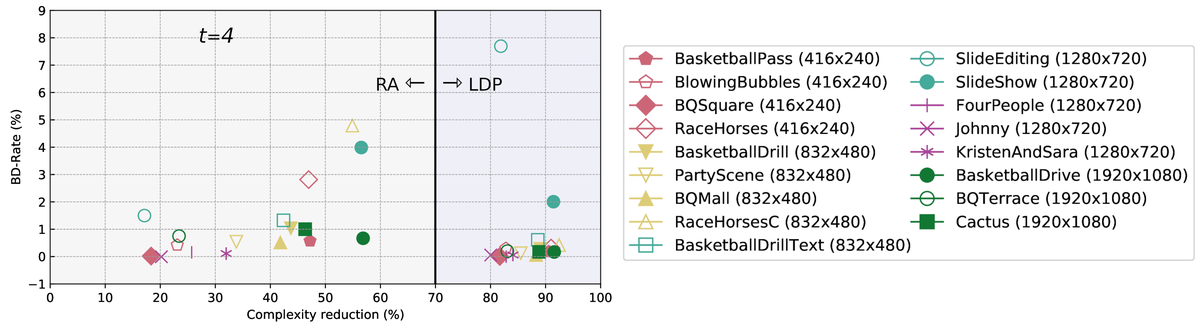

Figure 4: Experimental results of rate-based candidate elimination applied, with a fixed **threshold of t=4**, over TZS.

RA configuration results

When using **t=20** as our elimination threshold in RA, we

- approximate the Octagonal-axis pattern
- achieve BD-Rate $< 1\%$

Sequence	t=20		Octagonal-axis [6]	
	BD-Rate	ΔC	BD-Rate	ΔC
Cactus	0.12	28.8	0.02	26.2
BasketballDrill	0.14	23.4	0.01	22.2
BasketballDrillText	0.09	22.9	-0.04	21.2
SlideEditing	0.63	6.8	0.03	6.1
RaceHorses	0.35	25.1	0.02	22.5
SlideShow	0.88	42.5	-0.05	36.4
RaceHorsesC	0.70	33.9	0.10	30.2

LDP configuration results

A fixed threshold of **t=4** in LDP yields

- complexity reductions $> 80\%$
- negligible BD-Rate increases
  - only exceptions: screen-content sequences

Class	t=4		Octagonal-axis [6]	
	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

Conclusions

- In this work, we show that MV bitrate influences the efficacy of IME search patterns
- We propose an algorithm that can **reduce ME complexity by 86.69% at the cost of only an average 0.74% BD-Rate increase**
- Our results indicate that **the IME search can be drastically simplified** in some configurations

References

[1] F. Bossen.
Common HM test conditions and software reference configurations.
Document JCTVC-L1100, JCT-VC of ITU-T, Geneva, Jan 2013.

[2] F. Bossen, J. Boyce, X. Li, V. Seregin, and K. Sühring.
JVET common test conditions and software reference configurations for SDR video.
Document JVET-N1010, JVET of ITU-T, Geneva, Mar 2019.

[3] Cisco.
Cisco Visual Networking Index: Forecast and Trends, 2017–2022 White Paper, 2019.
Accessed: 20 Aug. 2019.

[4] R. Fan, Y. Zhang, and B. Li.
Motion Classification-Based Fast Motion Estimation for High-Efficiency Video Coding.
IEEE Transactions on Multimedia, 19(5):893–907, May 2017.

[5] P. Goncalves, G. Correa, M. Porto, B. Zatt, and L. Agostini.
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, Oct 2017.

[6] P. Goncalves, M. Porto, B. Zatt, L. Agostini, and G. Correa.
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, Apr 2018.

[7] I. Siqueira, G. Correa, and M. Grellert.
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, Feb 2020.

[8] G. J. Sullivan and T. Wiegand.
Rate-distortion optimization for video compression.
IEEE Signal Processing Magazine, 15(6):74–90, 1998.

[9] Y. Zhang, C. Zhang, and R. Fan.
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, Nov 2018.

Contact: baiocchi.gabriel@gmail.com
Associated code: gitlab.com/baioc/vtm/tree/rate-elimination

International Conference on Acoustics, Speech, and Signal Processing (ICASSP)
June 06-11, 2021 – Toronto, Canada