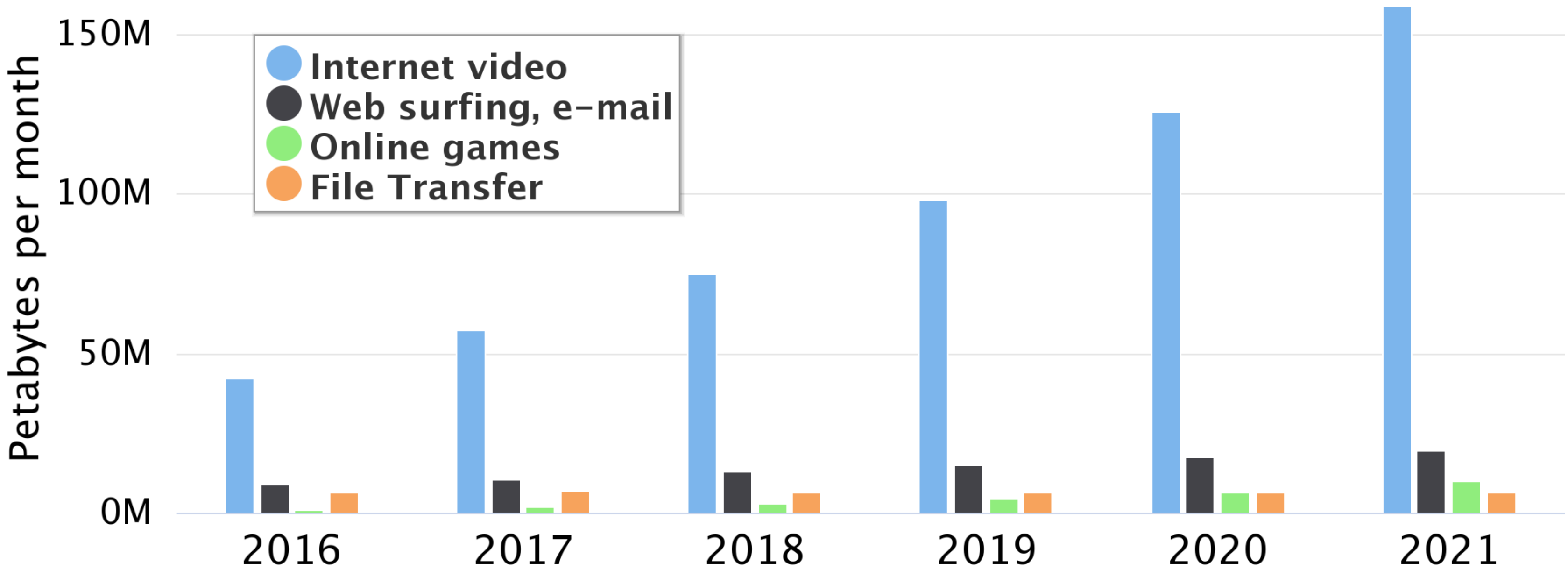
Iterative Machine-Learning-Based Method of Selecting Encoder Parameters for Speed-Bitrate Tradeoff

Sergey Zvezdakov, Alexey Solovyov, Dmitriy Vatolin

Video Group
CS MSU Graphics&Media Lab
Lomonosov Moscow State University

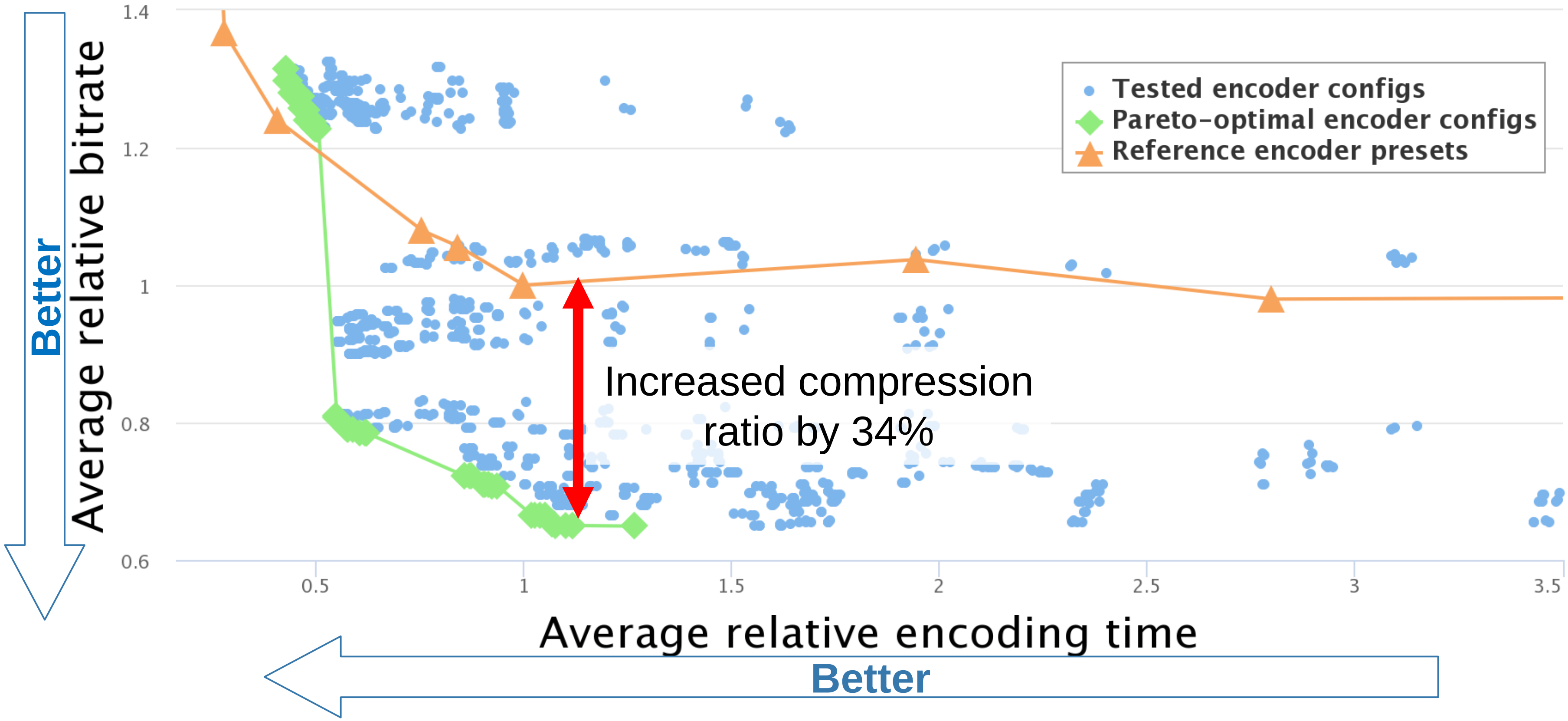
Motivation



- By 2022, video will account for 82% of all Internet traffic
- Modern video encoders are complex systems
  - Encoder x264 — 47 meaningful parameters
  - Affect the speed/quality of encoding in an unobvious way

According to the Cisco VNI Report 2018 and InterDigital Report 2020

Selecting encoder configuration to increase compression ratio



To encode a video that is **only 500 frames long (20 seconds)** with all possible x264 configs, it would take ~$\mathbf{2 * 10^{13}}$ **centuries of calculations (~488,000 ages of Earth)**

Problem formulation

Input:

- Tested video sequence V
- Testes video encoder C
- $Z_i = \{z_1, z_2, \dots, z_{M_i}\}$ — set of possible values for option i of codec C, $1 \le i \le N_C$
- $r = \{p_1, p_2, \dots, p_{N_C}\}$ — encoder's config, $p_i \in Z_i$
  The set of all possible configs r is referred to as P

Output:

- Set of Pareto optimal configs P_{pareto} for video sequence V and video codec C

Used datasets

- 1145 video sequences
  - almost every public dataset for testing video encoders
  - samples from 14Tb of videos downloaded from Vimeo.com
- Compute cluster of 253 identical computers
- 3 different video encoders
  - option lists available in the supplementary materials

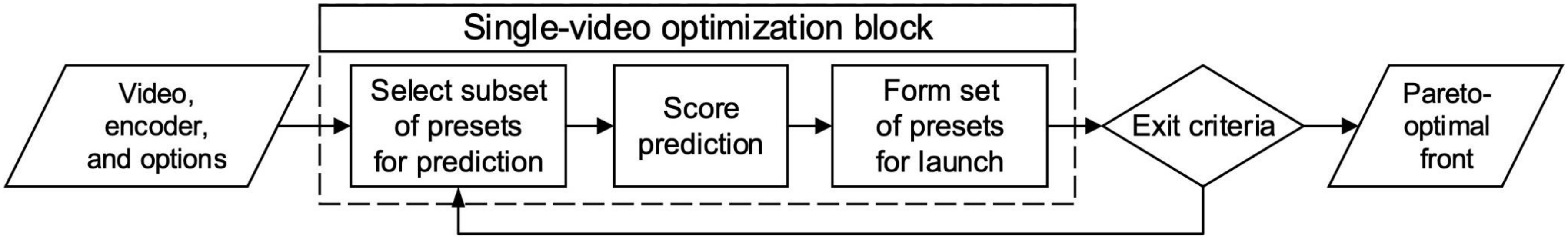
Encoder	Number of options	Option types			Videos tested	Unique configs tested
		Cat.	Integer	Float		
x264	47	18	22	7	1 145	297 829
x265	63	35	24	4	1 145	70 648
Commercial H.265	412	132	263	17	28	28 038

List of tested video encoders and their respective option statistics.

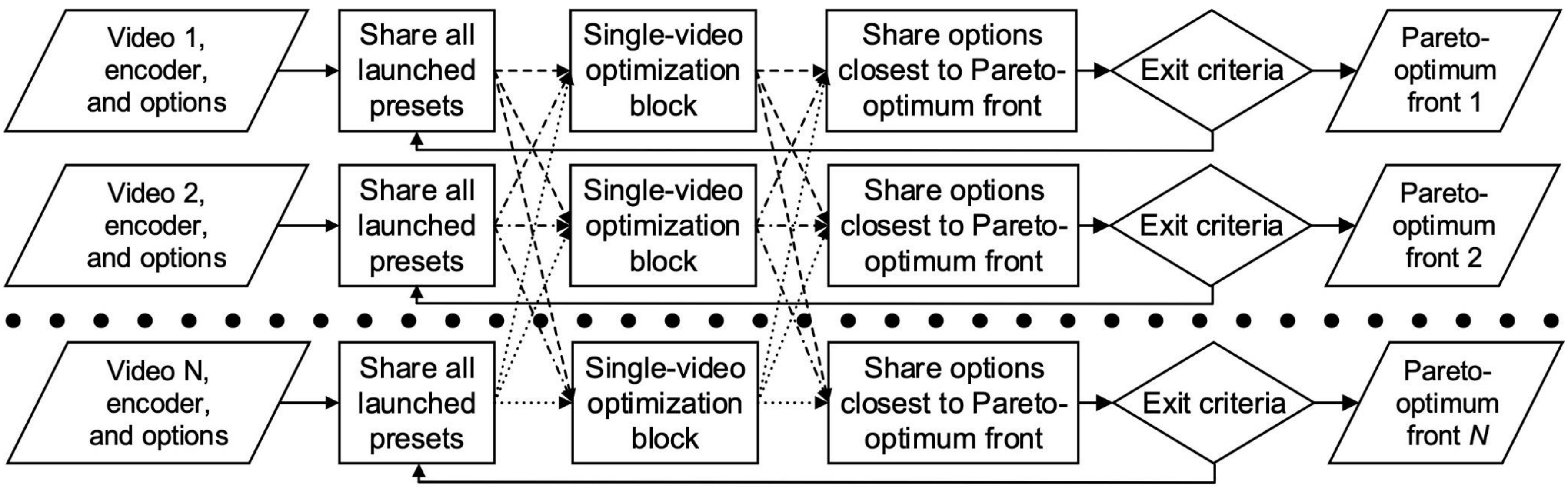
Information about video datasets and encoders options:
https://papers.evt.guru/dcc-ml-preset-miner/info.html

Scheme of the proposed method

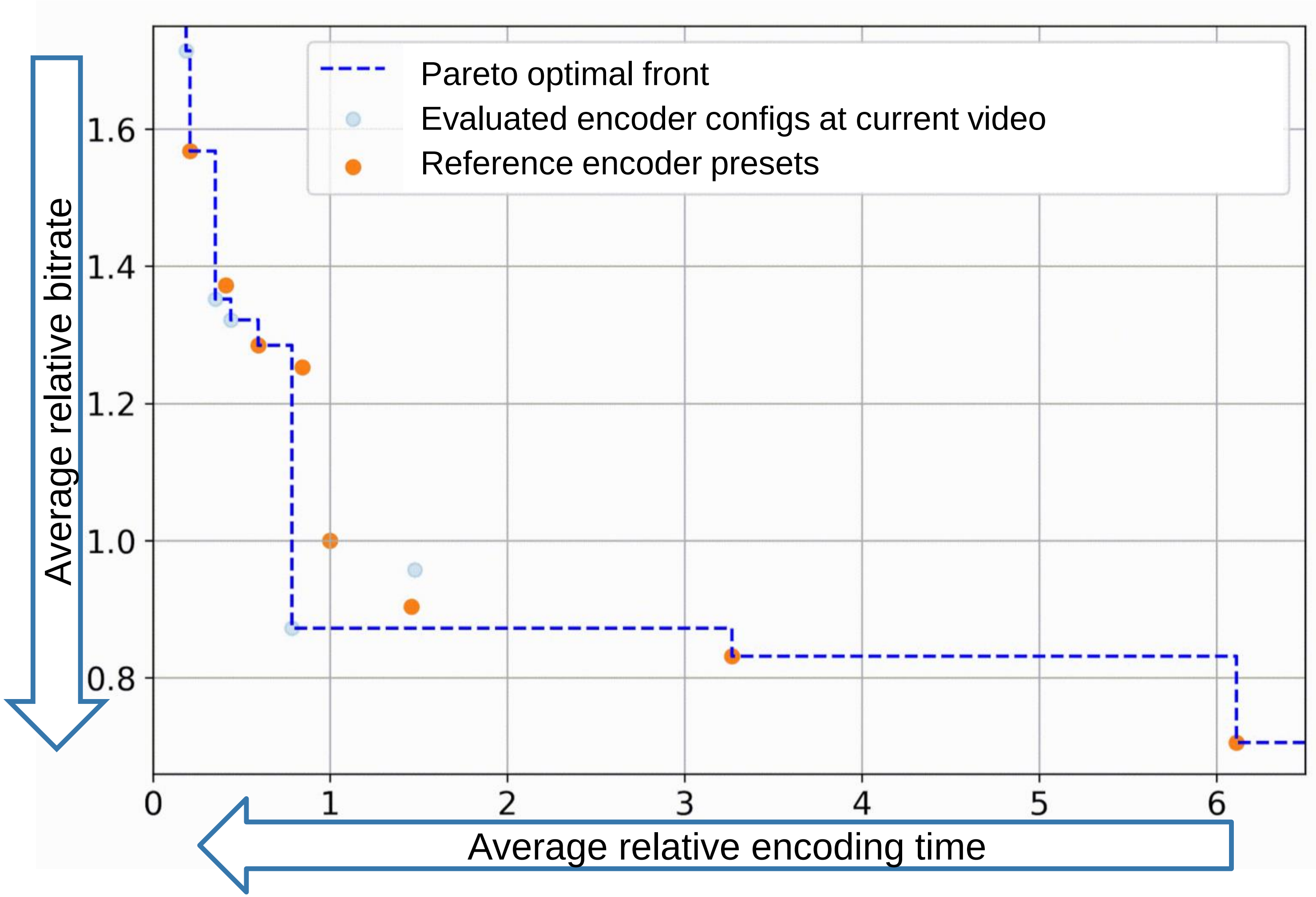
Basic scheme of the method (single video)



Proposed method scheme for working with multiple videos



Example



Examples for different videos and methods available here:
https://papers.evt.guru/dcc-ml-preset-miner/info.html

Comparison

Procedure and overall results

We compare proposed method with available MOO algorithms

- Codec x264 and its 4 reference presets
- 20 videos from 4 public video datasets
- Budget of 400 tested configs for each method on each video

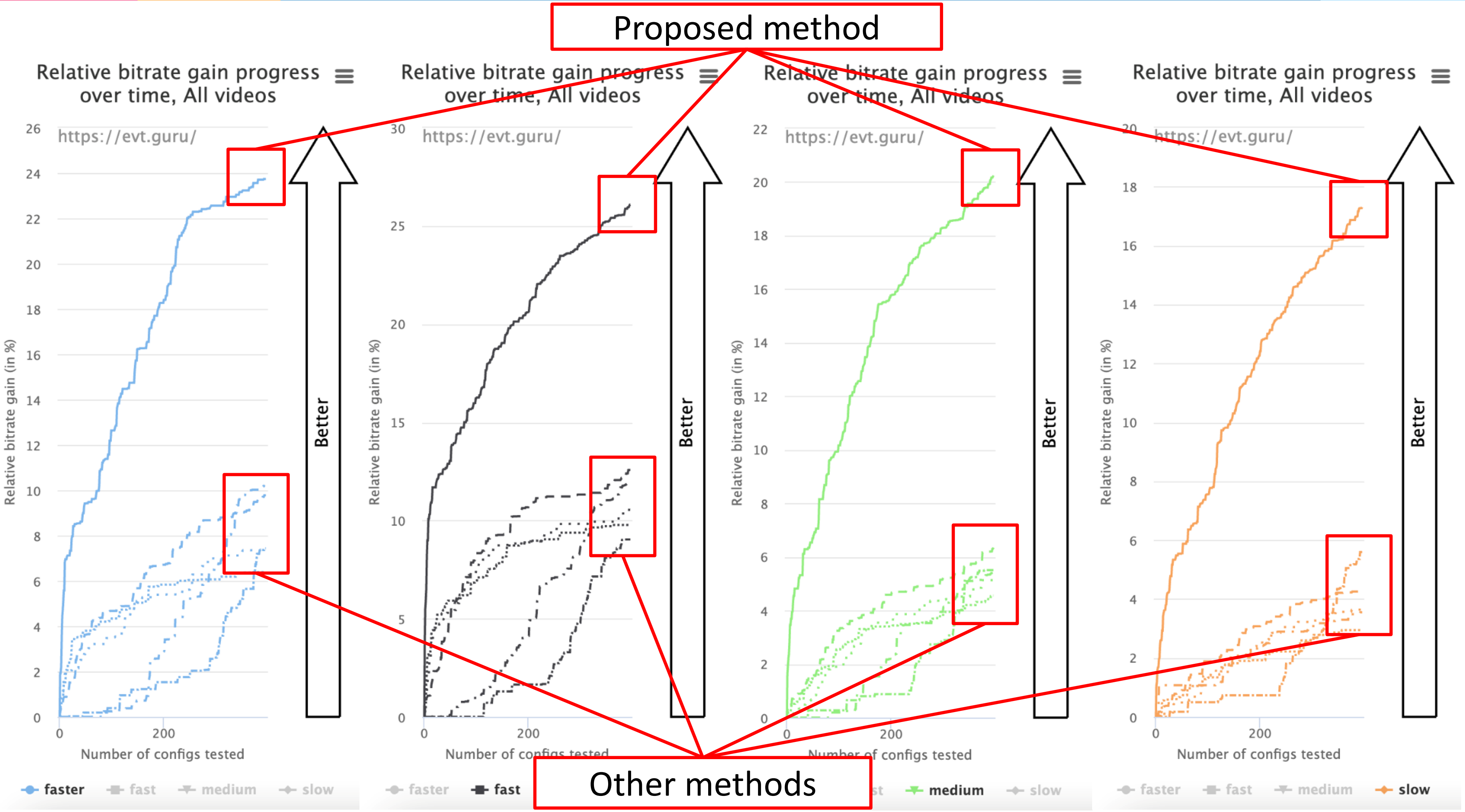
	faster		*fast*		*medium*		*slow*	
	Gain (%)	HV	Gain (%)	HV	Gain (%)	HV	Gain (%)	HV
NSGA-II [3]	7.4	0.007	10.6	0.014	5.1	0.008	2.6	0.014
ParEGO [13]	7.4	0.012	9.0	0.021	5.5	0.019	5.6	0.036
NSGA-III [4]	6.4	0.006	9.8	0.013	4.6	0.006	3.0	0.011
AGE-MOEA [6]	9.8	0.012	13.6	0.022	6.3	0.011	4.3	0.020
qEHVI [10]	10.3	0.017	12.0	0.031	5.4	0.017	3.5	0.022
Proposed	**23.7**	**0.026**	**26.1**	**0.066**	**20.2**	**0.061**	**17.3**	**0.123**

Comparison of the proposed method with multiobjective optimization algorithms

Detailed results available here:
https://papers.evt.guru/dcc-ml-preset-miner/info.html

Comparison

Progress rate



CS MSU Graphics&Media Lab (Video Group)
www.compression.ru/video

Charts available here:
https://papers.evt.guru/dcc-ml-preset-miner/info.html

Conclusion

- We propose a new approach for searching Pareto-optimal encoder configurations
  - It is independent of the video encoder's architecture and implementation
  - It outperforms all other MOO algorithms on our task
- Preset space study for x264 and x265 yielded new presets (with same quality and encoding time) which achieve:
  - 17–23% bitrate savings over four standard presets for independent video tuning
  - 3.6–5.6% bitrate savings over two standard presets for simultaneous video tuning