2020 Data Compression Conference

Machine-Learning-Based Method for Finding Optimal Video-Codec Configurations Using Physical Input-Video Features

GRAPHICS & MEDIA LAB
VIDEO GROUP

Roman Kazantsev, Sergey Zvezdakov, and Dmitriy Vatolin

Graphics & Media Lab, Lomonosov Moscow State University, Russia

{roman.kazantsev, sergey.zvezdakov, dmitriy.vatolin}@graphics.cs.msu.ru

Problem definition

One of the main video compression challenges is to configure a video codec so that a preset meets user's requirements for encoding time and video-quality loss. For a video v on presets P it can be formulated as a multicritireon-optimization problem:

$$(Q_v(p), T_v(p)) \to min, p \in P \qquad (*)$$

$Q_v(p), T_v(p)$ — the average bitrate/time required to encode a quality/bitrate unit.

Find: $P^*(v)$ — approximate solution of $(*)$, i.e. approximate Pareto-optimal set of configurations.

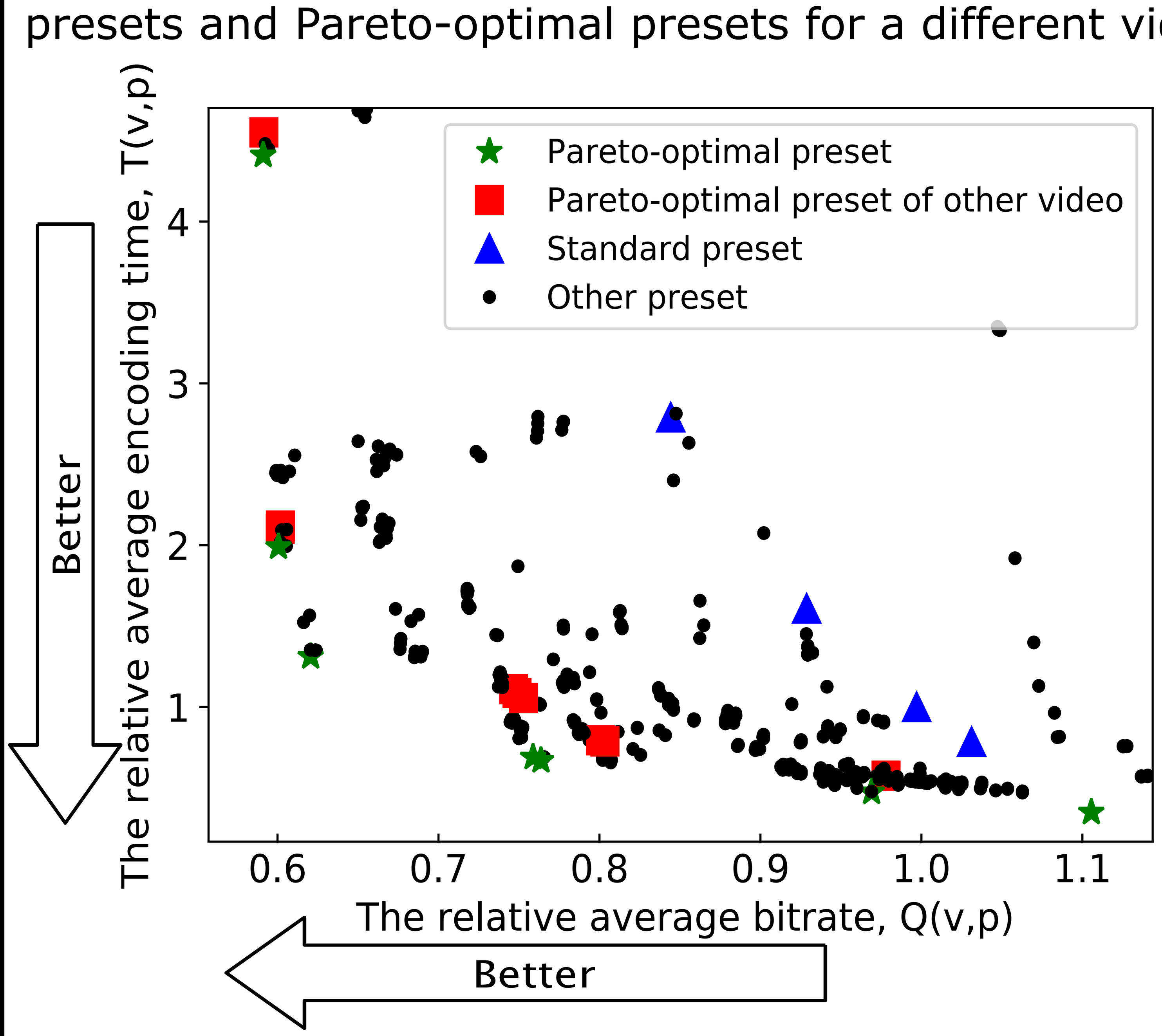
Dataset creation

For dataset creation we selected 355 videos from vimeo.com and 1306 presets of x264 video codec.

The following were computed for each video-preset pair:

- encoding time
- objective quality metric — SSIM
- a size of bitstream resulted by encoding

Example below demonstrates inefficiency of standard presets and Pareto-optimal presets for a different video:



Proposed method

Training

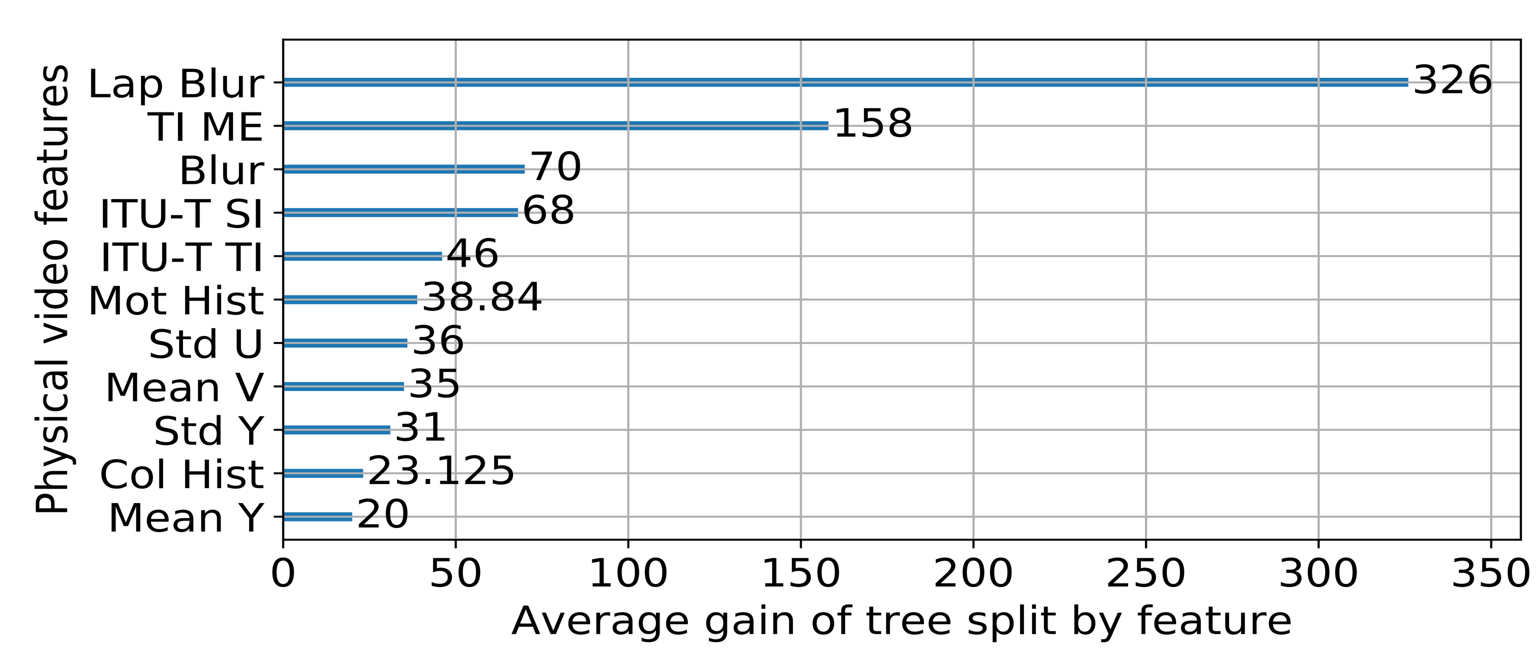
1. Cluster videos according to similarity of Pareto-frontier structures. Four clusters were obtained
2. Assign to each cluster the Pareto-optimal set of some video from this cluster
3. Train a model that predicts a cluster using the physical video features [4]

Inference

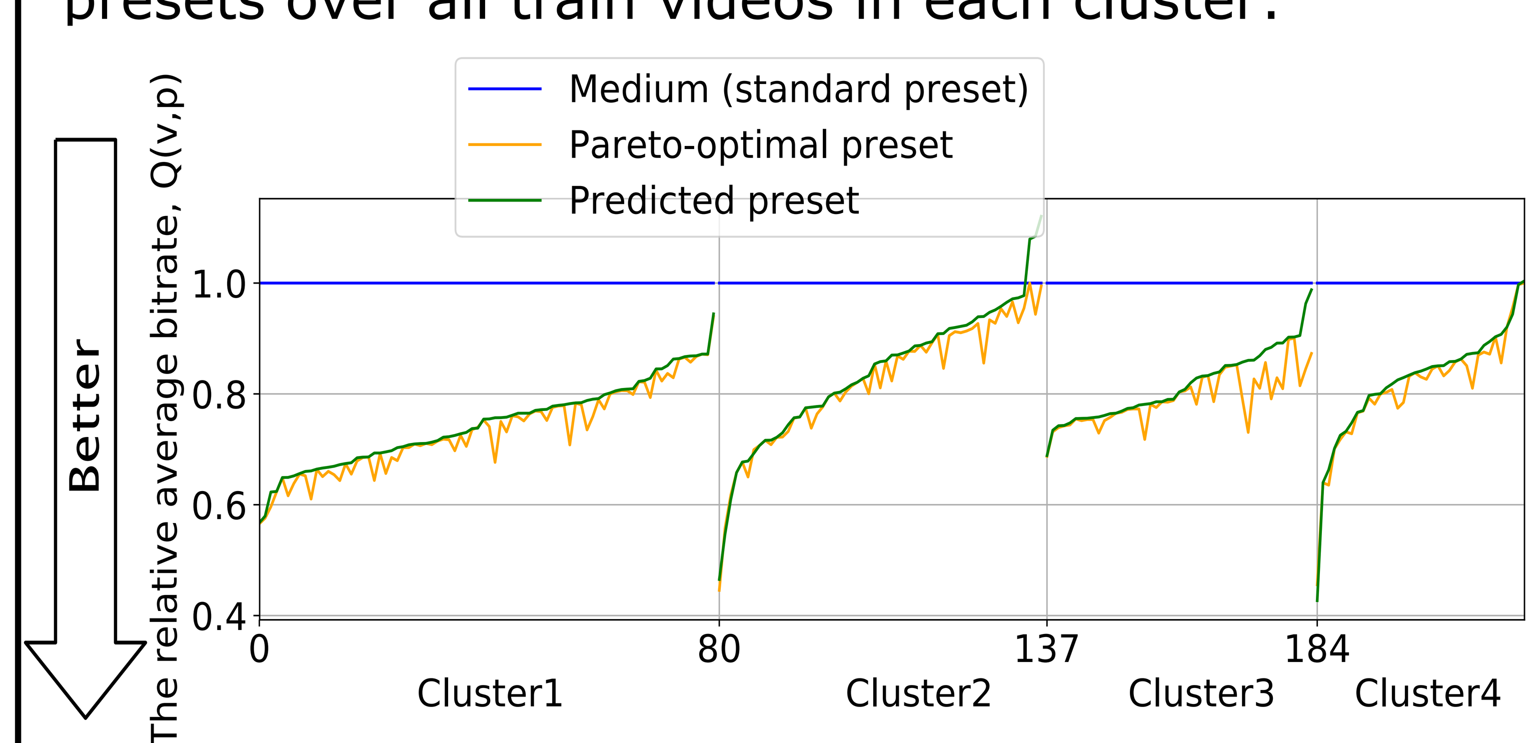
1. Compute the physical features for input video
2. Predict a cluster using the model and output Pareto-optimal set assigned to the predicted cluster

Results

Importance of physical video features in trained model:



Bitrates delivered using optimal, predicted and standard presets over all train videos in each cluster:



Results cont.

Average bitrate savings [%] of the predicted presets versus standard presets and execution time obtained using different methods:

Method	Faster	Fast	Medium	Slow	Slower	Veryslow	Placebo	Time, sec.
NSGA-II [1]	15.9	30.2	29.7	34.9	32.2	29.0	28.3	13686.4
Popov's [2]	8.0	29.0	28.4	34.9	32.2	28.3	29.0	10039.7
Zvezdakov's [3]	11.0	30.2	29.8	34.9	32.2	28.7	29.4	7705.2
Ours	15.8	21.3	21.8	27.7	24.9	9.7	10.5	**735.5**

Bitrate savings [%] obtained using the predicted presets versus the standard presets on JVET videos:

Sequence	Faster	Fast	Medium	Slow	Slower	Veryslow	Placebo
Cactus	6.0	16.5	17.5	15.0	13.3	8.0	15.7
DugsAndLegs	17.2	27.9	26.5	21.0	15.4	1.3	11.5
KristenAndSara	4.5	12.3	16.0	24.7	25.1	0.0	9.0
ParkScene	12.6	17.5	29.0	22.8	19.0	14.1	16.3
PeopleOnStreet	16.8	21.9	27.6	33.7	31.4	20.2	29.8
Average	**11.4**	**19.2**	**23.3**	**23.4**	**20.8**	**8.7**	**16.4**

Conclusions

1. The proposed method finds presets that provide 9-20% bitrate savings against x264 standard presets
2. The method slightly looses to existing solutions in bitrate saving, however, it is faster by 10 times than existing solutions
3. It can be applied to other video codecs and standards
4. A good dataset creation for video codec modeling is hard and time-consuming process
5. Acutance metric — *Lap Blur* and temporal complexity — *TI ME* are the most relevant physical video features

References

[1] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182-197, August 2002.

[2] V. Popov, "Automatic method of choosing pareto optimal video codec's parameters," M.S. thesis, Lomonosov Moscow State University, 2009.

[3] S. Zvezdakov and D. Vatolin, "Building a x264 video codec model," in *Innovative technologies in cinema and education: IV International Symposium*, Moscow, Russia, 2017, pp. 56-65, VGIK Moscow.

[4] I. Brailovskiy and N. Solomeshch, "Quality modelling for videocoding," *Information Technologies*, vol. 1, no. 1, pp. 42-48, 2012.

Acknowledgement: This work was partially supported by Russian Foundation for Basic Research under Grant 19-01-00785a