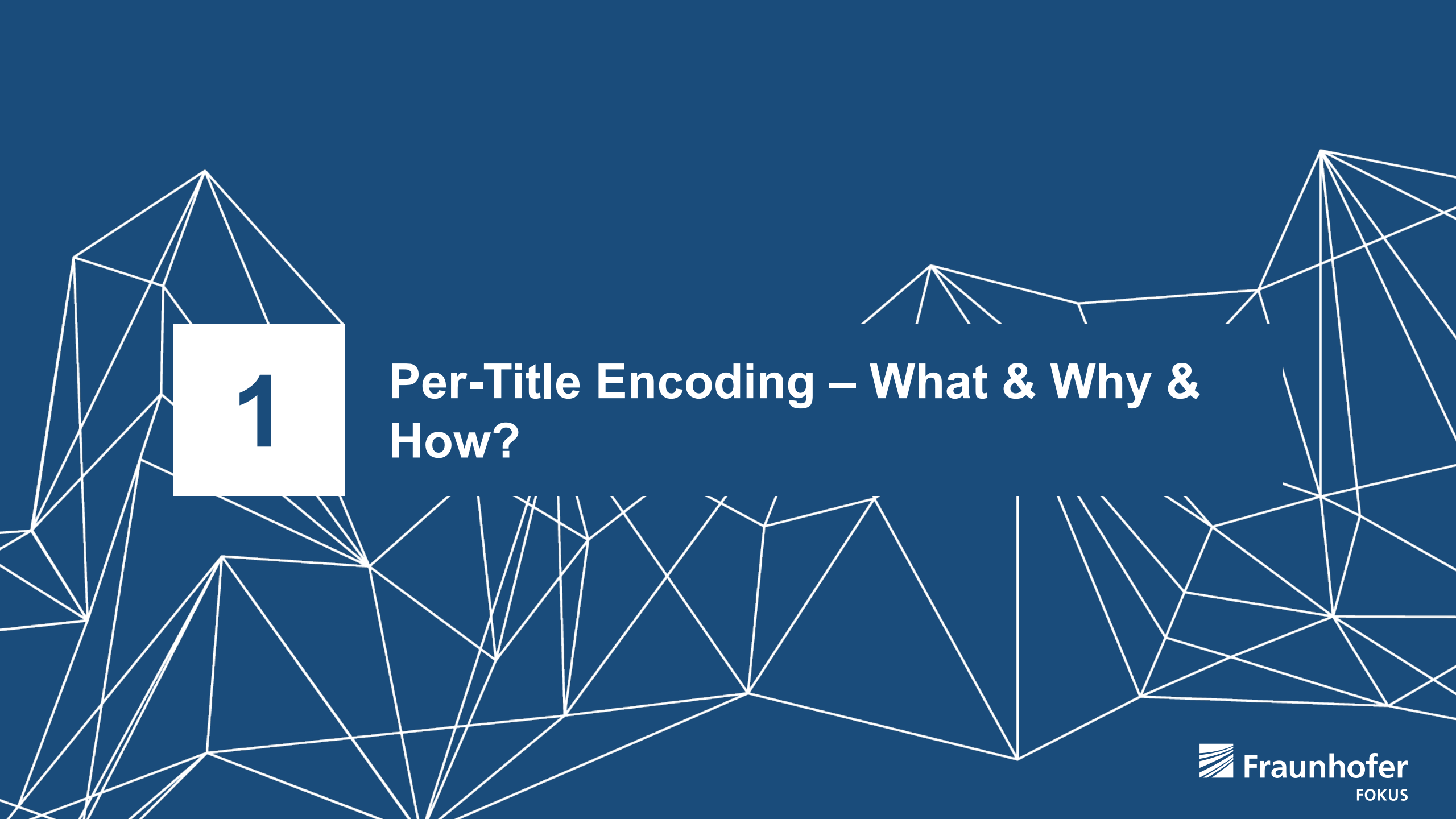


Deep Encode: Machine Learning for Per-Title Encoding

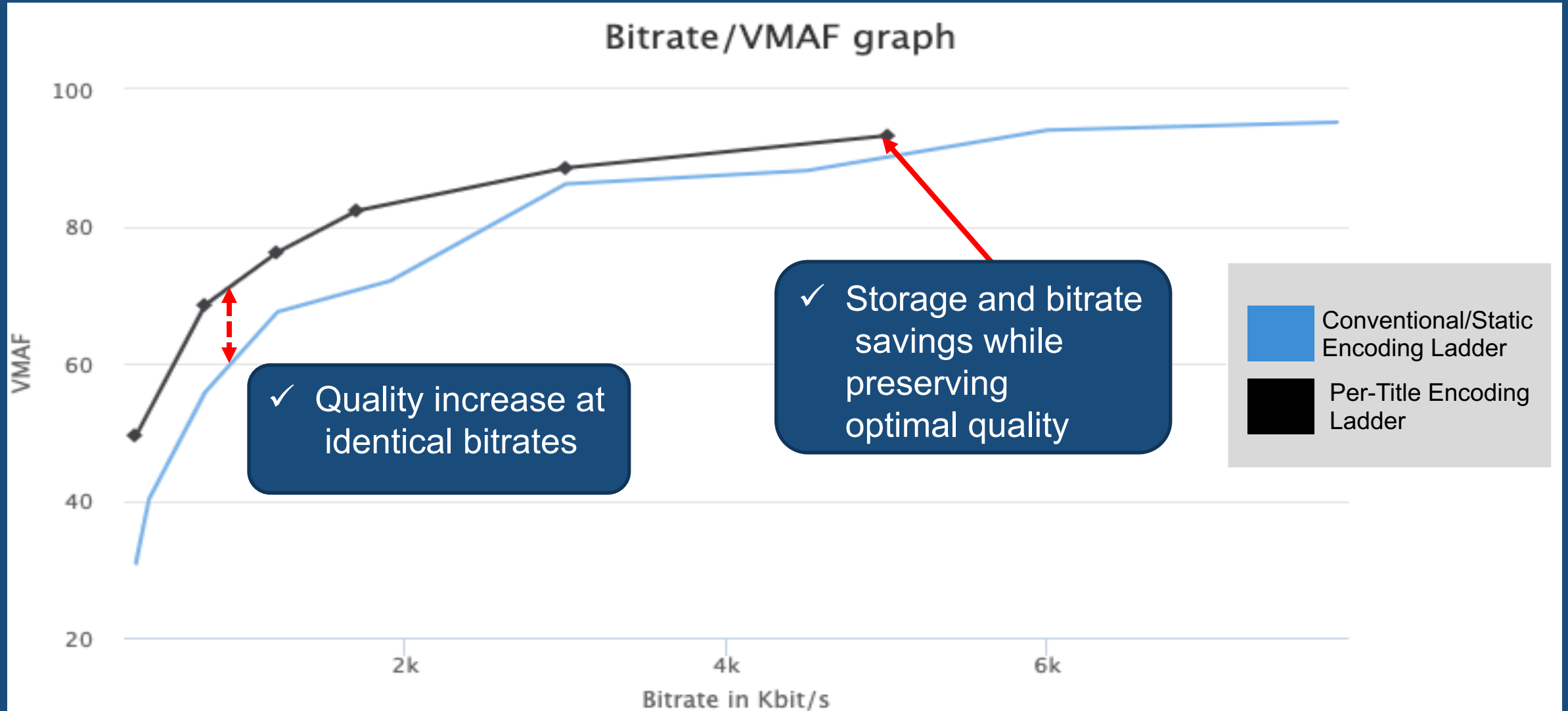
Daniel Silhavy | IBC20 |



1

Per-Title Encoding – What & Why & How?

Benefits of Per-Title Encoding



Benefits of Per-Title Encoding: the Numbers

	Average values			
	Bitrate (kbit/s)	VMAF	PSNR (dB)	Storage (MB)
Conventional	7648.18	94.92	44.37	1397.7
Per-Title	4941.75	93.06	42.41	675.2
Difference Abs	+2706.43	-1.86	-1.96	+722.5 MB
Difference (%)	+36%	-1%	-4%	+52%

* Based on a streaming session with a 100Mbit/s connection in dash.js 3.1.1

How does Per-Title Encoding work?

Test Encodes

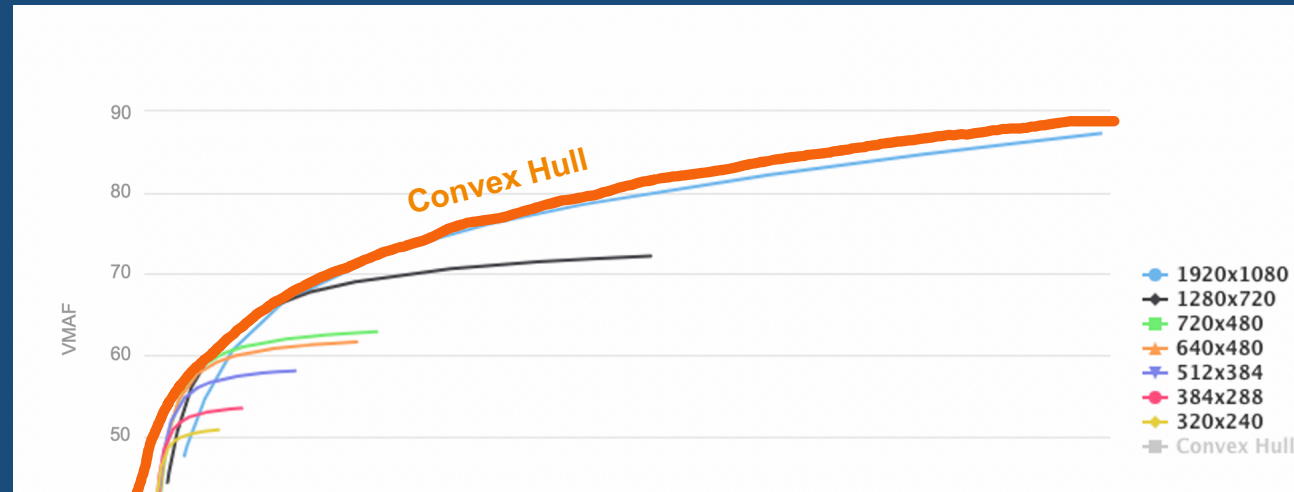
- Perform test encodings with different settings and calculate corresponding VMAF values.

Convex Hull Estimation

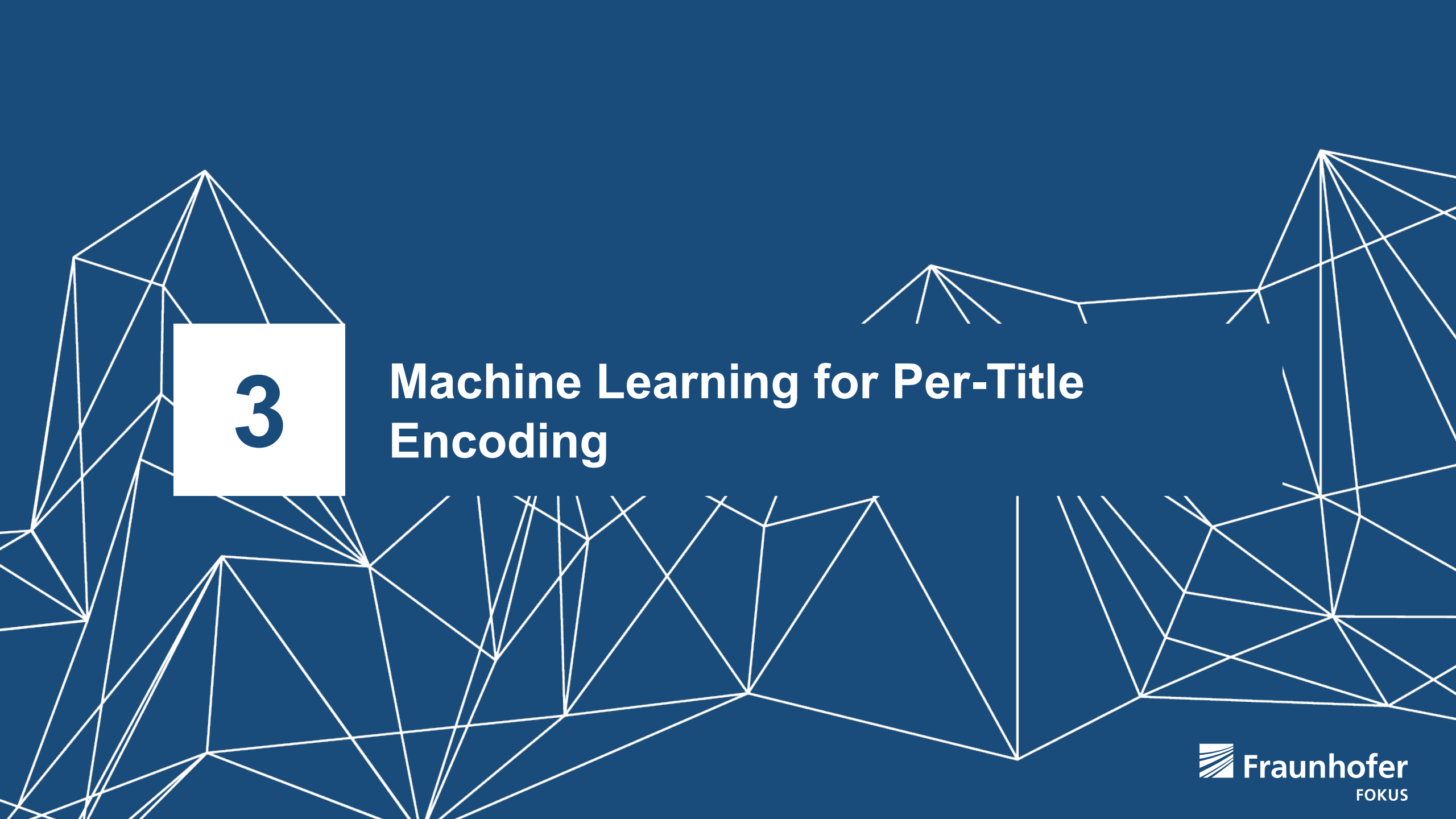
- Select bitrate-resolution pairs that are close to the convex hull.

Production encoding

- Perform the production encoding using the optimal encoding ladder.



A large amount of test encodes is required to derive a sufficient amount of data points.



3

Machine Learning for Per-Title Encoding

How to avoid the computationally heavy test encodes

ML-based predictions

- Predict mandatory Bitrate / VMAF pairs using machine learning.

Convex Hull

- Select bitrate-resolution pairs that are close to the convex hull

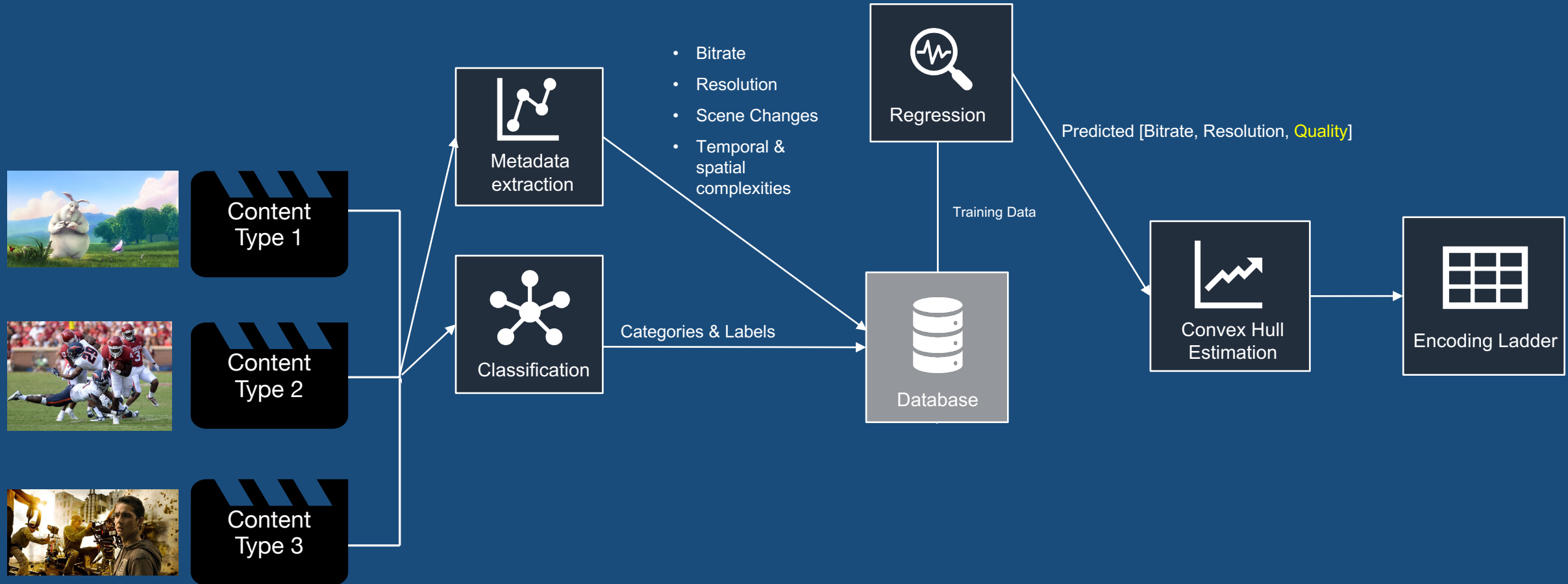
Production encoding

- Perform production CBR / Two-pass constrained VBR encoding using the optimal encoding ladder



Use ML-based predictions to avoid test encodes and *still* derive a sufficient amount data points.

Deep Encode – ML for Video Metric Prediction



Deep Encode: Hands-on UI

Upload Video

Sample Videos

Clip History

↑

Drag and Drop video file here or click the button to upload.

Maximum allowed file size is 150 MB.

Select Video

Legal Information

Data Protection

© Fraunhofer FOKUS

Upload Video

Sample Videos

Clip History

YOUR UPLOAD

output_video.mp4

Length00:00:05

Bitrate672 kb/s

Resolution1920x1080

Framerate30 FPS

Size481.74 KB

CodecH.264

Video Analysis

Characteristics Calculation

Video Content Classification

Scene Change Detection

SI / TI Score Calculation

Started a few seconds ago

Started a few seconds ago

Started a few seconds ago

Started a few seconds ago

Started a few seconds ago

100 %

13 %

100 %

100 %

0 %

All jobs done!

1/1 jobs running

All jobs done!

All jobs done!

1/1 jobs running

Updated: 10:05:48

Legal Information

Data Protection

© Fraunhofer FOKUS

Upload Video

Sample Videos

Clip History

Settings

Upload New Video

YOUR UPLOAD

output_video.mp4

Length00:00:05

Bitrate672 kb/s

Resolution1920x1080

Framerate30 FPS

Size481.74 KB

CodecH.264

Customize your desired video settings like your minimum / maximum bitrate or VMAF preference. These settings will be used during the predictions and are applied to the encodes, which can be created in the final steps of the process. In case you don't want to customize anything, simply use the default settings.

Video Bitrate (kb/s)

Range 1000 - 4000 kb/s

Factor 1.25 - 1.75

VMAF Score

Range 40 - 93 points

Distance 6 - 18 points

Tasks

☒ Scene Change Detection

☒ Video Content Classification

☒ Video Characteristics Extraction

☒ SI/TI Score Calculation

Start tasks

Restore default settings

Legal Information

Data Protection

© Fraunhofer FOKUS

Upload Video

Sample Videos

Clip History

Results

YOUR UPLOAD

part1_cut_hb.mp4

Length00:03:01

Bitrate452 kb/s

Resolution1920x1080

Framerate30 FPS

Size9.76 MB

CodecH.264

Predicted Encoding Settings

Final Encoding Settings

Bitrate-Quality Comparison

VMAF

145 kb/s1101 kb/s2058 kb/s3015 kb/s3972 kb/s4829 kb/s5886 kb/s6843 kb/s7800 kb/s

Apple H264 Encoding LadderFAMUIM Predicted Encoding Ladder

Reference Ladder

Apple H264 Encoding Ladder

#	Resolution	Bitrate (kb/s)	VMAF (predicted)
1	416 x 234	145	12
2	640 x 360	365	23
3	768 x 432	730	27
4	768 x 432	1100	45
5	960 x 540	2000	59
6	1280 x 720	3000	63
7	1280 x 720	4500	85
8	1920 x 1080	6000	94
9	1920 x 1080	7800	98

Predicted Encoding Ladder

FAMUIM DE - Logarithmic

#	Resolution	Bitrate (kb/s)	VMAF (predicted)
1	640 x 360	303	46
2	320 x 240	403	52
3	3840 x 2160	700	70
4	3840 x 2160	1423	88

Constraints

Bitrate min 1000 kb/s, max 4000 kb/s, Factor: 1.25 - 1.75

VMAF min 40, max 93, Stepsize: 6 - 18

The predicted values above follow the constraints you chose during the upload of your video. You can change these settings at any time. Once you are satisfied with the predicted encoding ladder, create the encodes by clicking on the button below.

Show Settings

Create 9 Encodes

Comparison

Bitrate Ladder	Available Representations	Average Bitrate (kb/s)	Average VMAF	Storage Size
Apple H264 Encoding Ladder	9	2849	56	35.41 MB
FAMUIM DE - Logarithmic	4	~2142 kb/s (75%)	707	+8 (+14%) 64 ~31.50 MB (89%) 3.91 MB



4

Summary, Outlook & Next Steps

Deep Encode: Towards Context-Aware Encoding



Conventional Static Encoding

- Same encoding ladder for all types of content
 - Increased storage and delivery costs
 - “Waste” of quality
 - Lack of optimization for complex content

Conventional Per-Title Encoding

- Computationally heavy test encodes
- No dynamic reaction to complex scenes within a movie

Deep Encode

- ✓ No computationally heavy test encodes
- ✓ Metadata extraction and AI-based image processing for content analysis
 - Content categorization and labeling
 - Automatic scene detection
 - Metadata extraction
- ✓ Deep Learning for optimal encoding ladders
 - Prediction of [PSNR|VMAF, Bitrate] pairs
 - Dynamic prediction of the optimal encoding ladder
- ✓ Enhancements
 - Live-stream support
 - Per-scene and context-aware Encoding

Thank you!



Daniel Silhavy

Project Manager Future Applications and Media (FAME)

daniel.silhavy@fokus.fraunhofer.de

Fraunhofer FOKUS

Berlin, Germany

FAME Video Development Blog: <https://websites.fraunhofer.de/video-dev/>