Fraunhofer FOKUS
Institut für Offene Kommunikationssysteme

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

Deep Encode:
Machine Learning for Per-Title Encoding

Daniel Silhavy| IBC20|

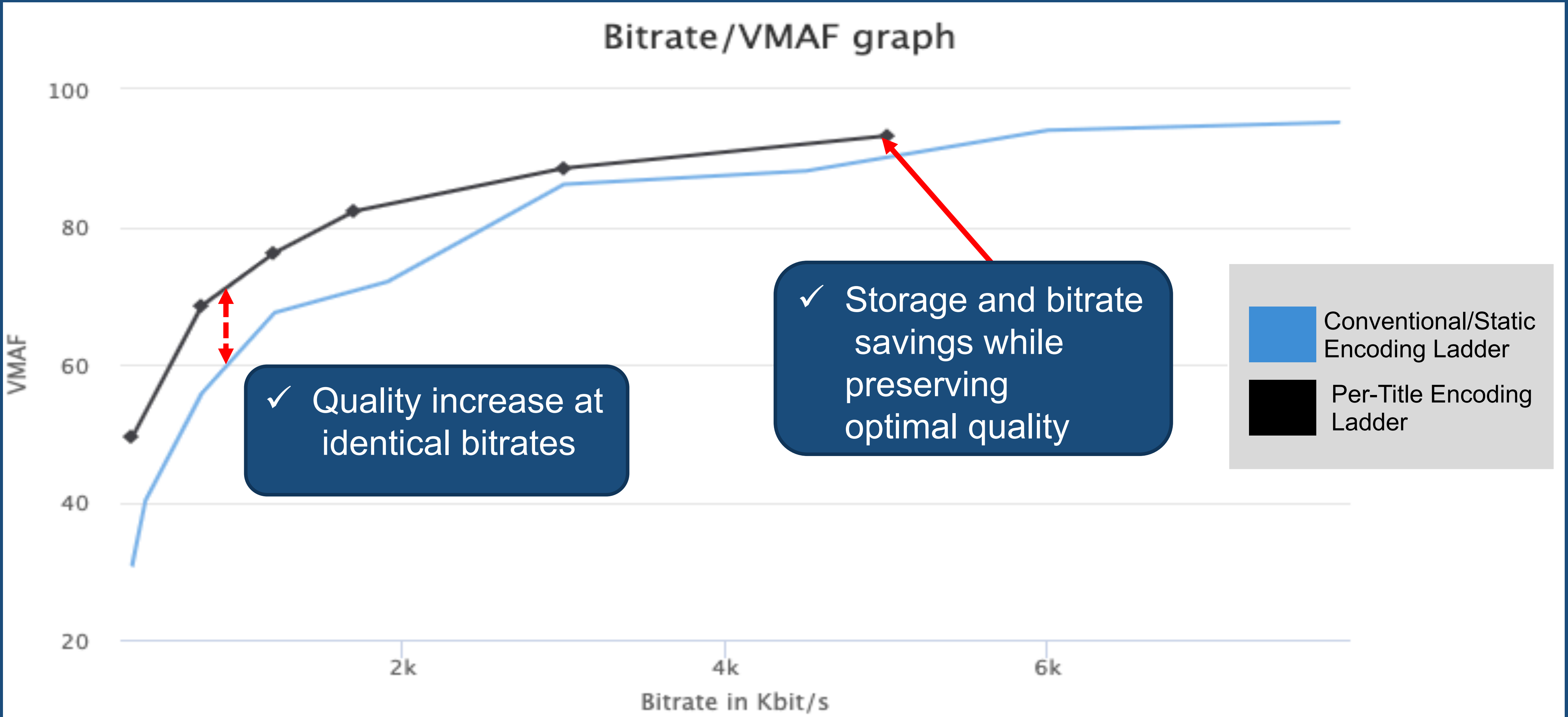
Fraunhofer FOKUS

1

Per-Title Encoding – What & Why & How?

Fraunhofer
FOKUS

Benefits of Per-Title Encoding



Fraunhofer FOKUS

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

Fraunhofer FOKUS

How does Per-Title Encoding work?

Test Encodes

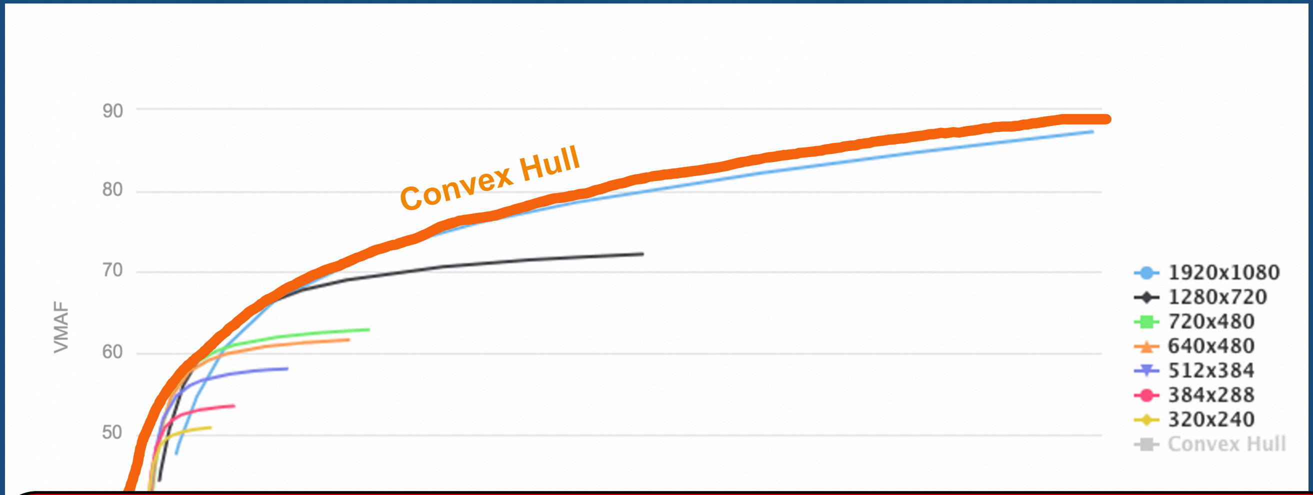
- Perform test encodings with different settings and calculate corresponding VMAF values.

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

Fraunhofer FOKUS

How to avoid the computationally heavy test encodes

ML-based predictions

- Predict mandatory Bitrate / VMAF pairs using machine learning.

Convex Hull

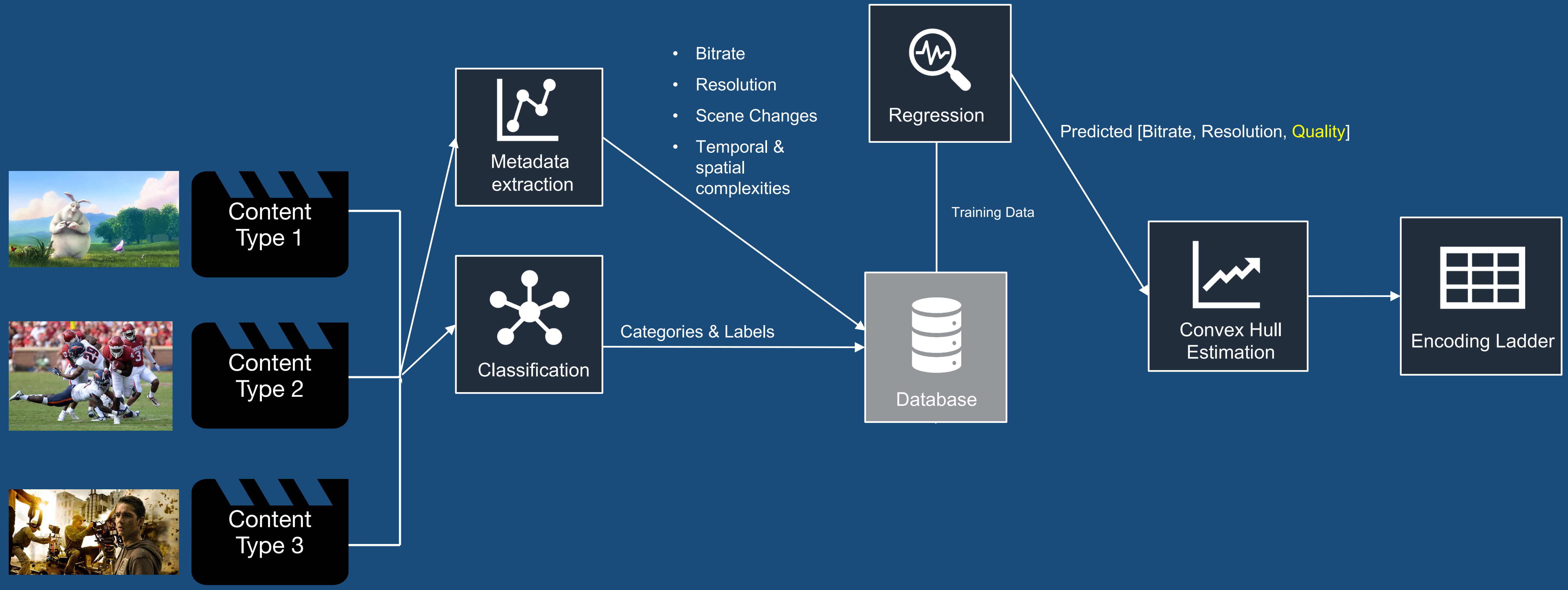
- Select bitrate-resolution pairs that are close to the convex hull

Production encoding

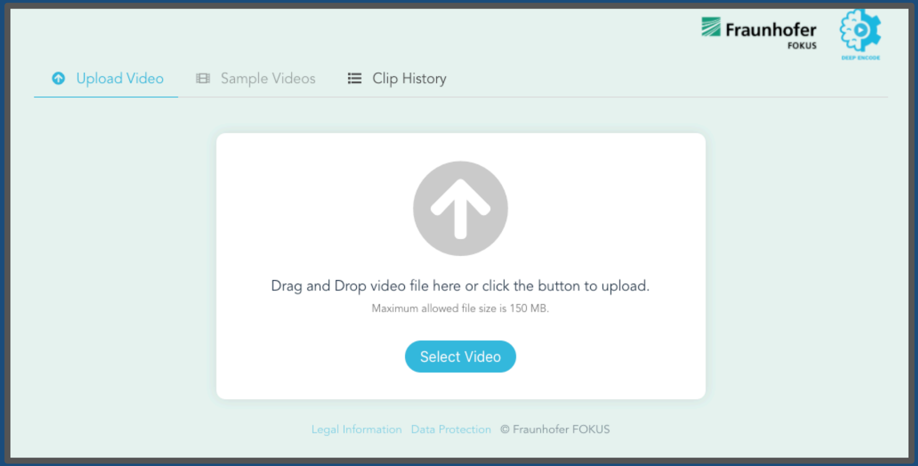
- Perform production CBR / Two-pass constrained VBR encoding using the optimal encoding ladder

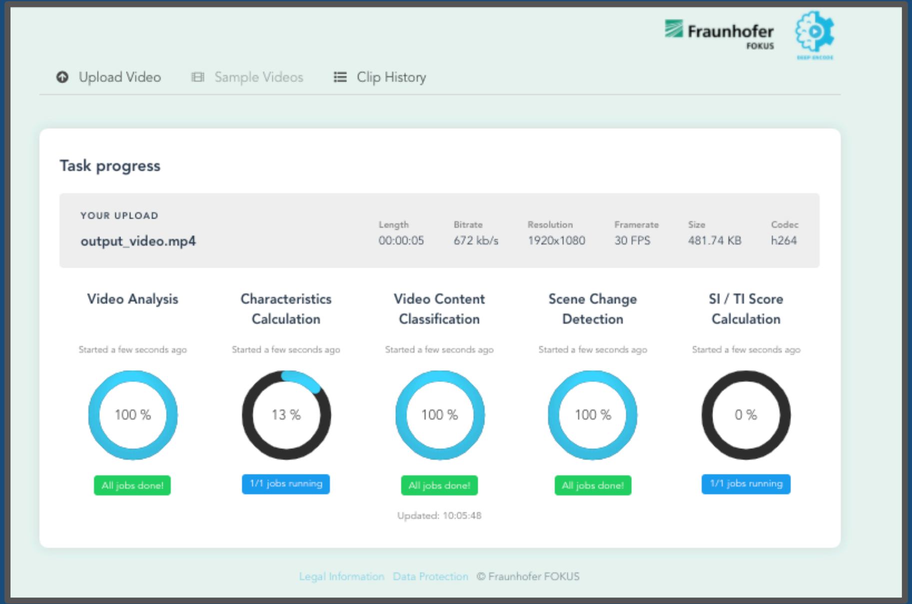
! Use ML-based predictions to avoid test encodes and *still* derive a sufficient amount data points.

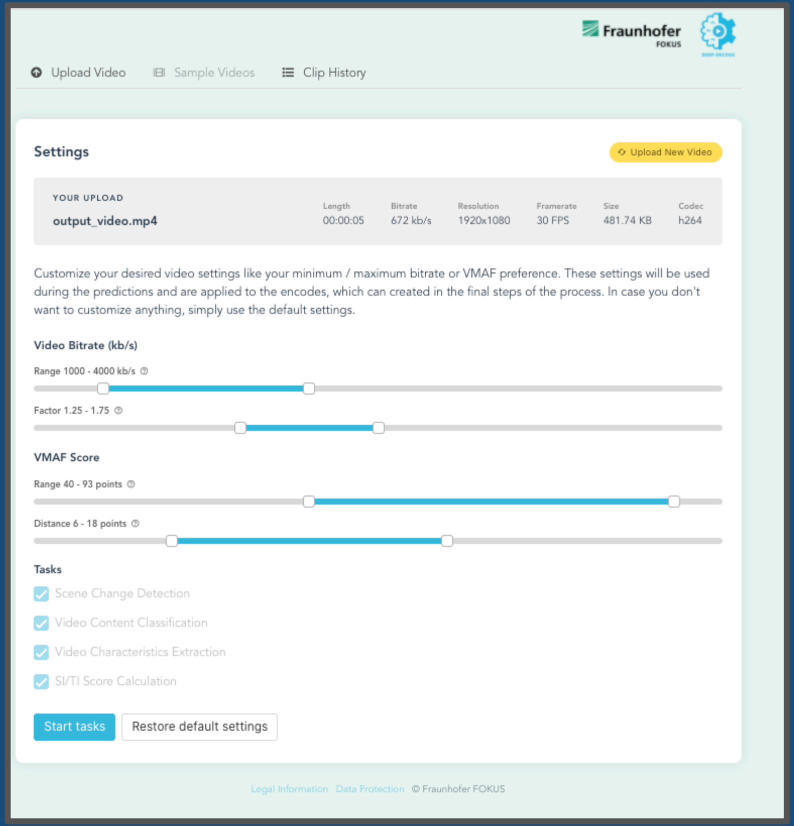
Deep Encode – ML for Video Metric Prediction

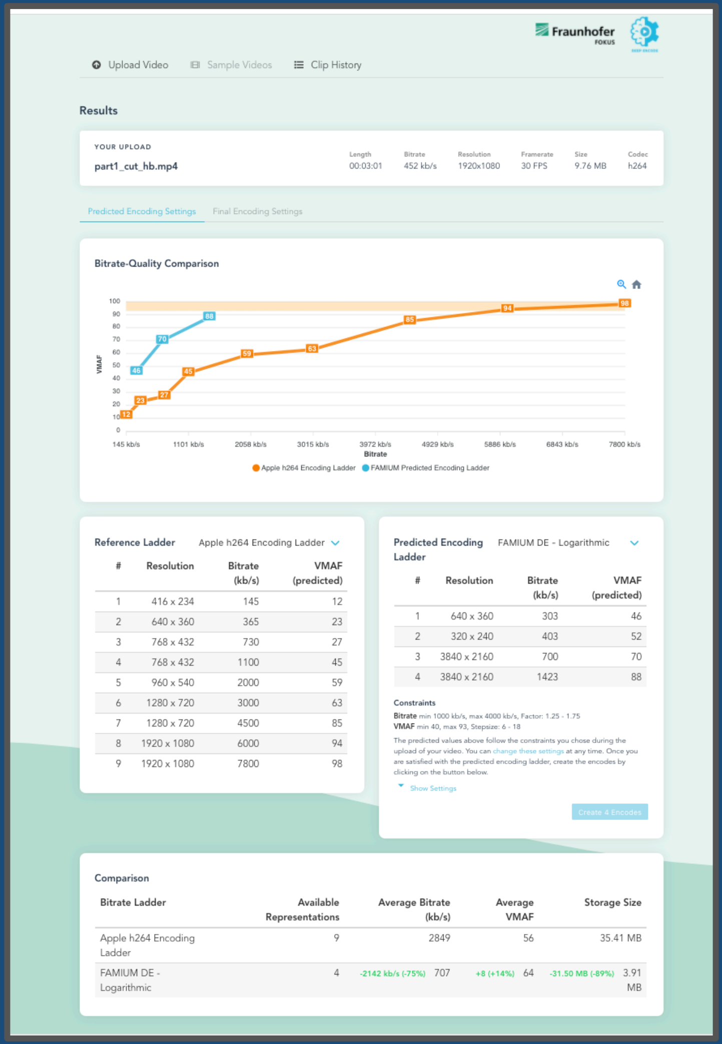


Deep Encode: Hands-on UI









4 Summary, Outlook & Next Steps

Fraunhofer FOKUS

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

Fraunhofer FOKUS

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/

Fraunhofer
FOKUS