Deep Encode:
Machine Learning for Per-Title Encoding

Daniel Silhavy | IBC20
Per-Title Encoding – What & Why & How?
Benefits of Per-Title Encoding

- Quality increase at identical bitrates
- Storage and bitrate savings while preserving optimal quality

Bitrate/VMAF graph

- Conventional/Static Encoding Ladder
- Per-Title Encoding Ladder
Benefits of Per-Title Encoding: the Numbers

<table>
<thead>
<tr>
<th></th>
<th>Average values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bitrate (kbit/s)</td>
</tr>
<tr>
<td>Conventional</td>
<td>7648.18</td>
</tr>
<tr>
<td>Per-Title</td>
<td>4941.75</td>
</tr>
<tr>
<td>Difference</td>
<td></td>
</tr>
<tr>
<td>Abs</td>
<td>+2706.43</td>
</tr>
<tr>
<td>(%)</td>
<td>+36%</td>
</tr>
</tbody>
</table>

* 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.
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

- **Content Type 1**
- **Content Type 2**
- **Content Type 3**

**Metadata extraction**
- Bitrate
- Resolution
- Scene Changes
- Temporal & spatial complexities

**Classification**

**Regression**
Predicted [Bitrate, Resolution, Quality]

**Database**
Training Data

**Convex Hull Estimation**

**Encoding Ladder**
Deep Encode: Hands-on UI

Customers can customize their desired video settings like the minimum frame rate to transmit or any VMF preferences. 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 150 - 4000 kbps
- Default 1500 kbps

VMAF Score
- Range 0 - 1.0
- Default 0.5

Tools
- Score Change Detection
- Video Content Classification
- Video Duplicates Detection
- VMF Score Calculation

Deep Encode – Machine Learning for Per-Title Encoding
Summary, Outlook & Next Steps
Deep Encode: Towards Context-Aware Encoding

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

**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

**Per-Scene & Per-Title Encoding**

**Conventional Per-Title Encoding**
- Computationally heavy test encodes
- No dynamic reaction to complex scenes within a movie

Fraunhofer FOKUS

Deep Encode – Machine Learning for Per-Title Encoding

12
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/