Impact of Spatial and Temporal Information on Video Quality and Compressibility

Werner Robitza[1,2], Rakesh Rao Ramachandra Rao[2], Steve Göring[2], Alexander Raake[2]

[1] AVEQ GmbH, Vienna, Austria
[2] Ilmenau University of Technology, Ilmenau, Germany

Video Quality Experts Group — 7-11 June, 2021

Intro

— This presentation is a preview of an upcoming paper at QoMEX 2021

— Online from June 14th-17th

— https://qomex2021.itec.aau.at/

Spatial Information / Temporal Information

About SI and TI:

— Defined in ITU-T Rec. P.910

— Classify spatiotemporal complexity of video sequences

— **SI:** Standard deviation of Sobel-filtered image

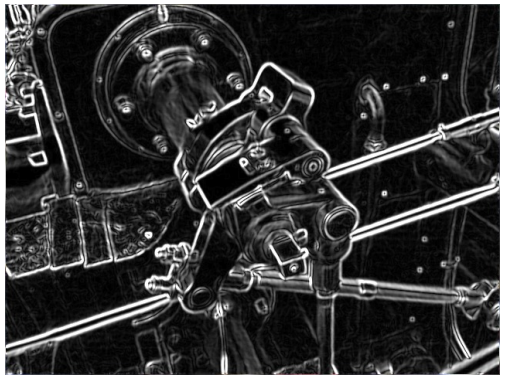
— **TI:** A basic motion difference feature for adjacent frames

This contribution:

— Related to ongoing VQEG NORM project

— SI/TI frequently used for classifying sources for video quality tests

— **How well can SI/TI be used to gauge compressibility of a video?**



By Simpsons contributor, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=8904663

This Contribution

- Related to ongoing VQEG NORM project with different activities
  - Usage for material with > 8 bit
  - HDR
  - …
- In this contribution, however, focus on:
  - SI/TI are frequently used for classifying sources for video quality tests
  - Choosing the sources is important for codec development, subjective testing, etc.
- **How well can SI/TI be used to gauge compressibility of a video?**

Working Hypothesis

Videos with higher SI/TI
should be harder to compress

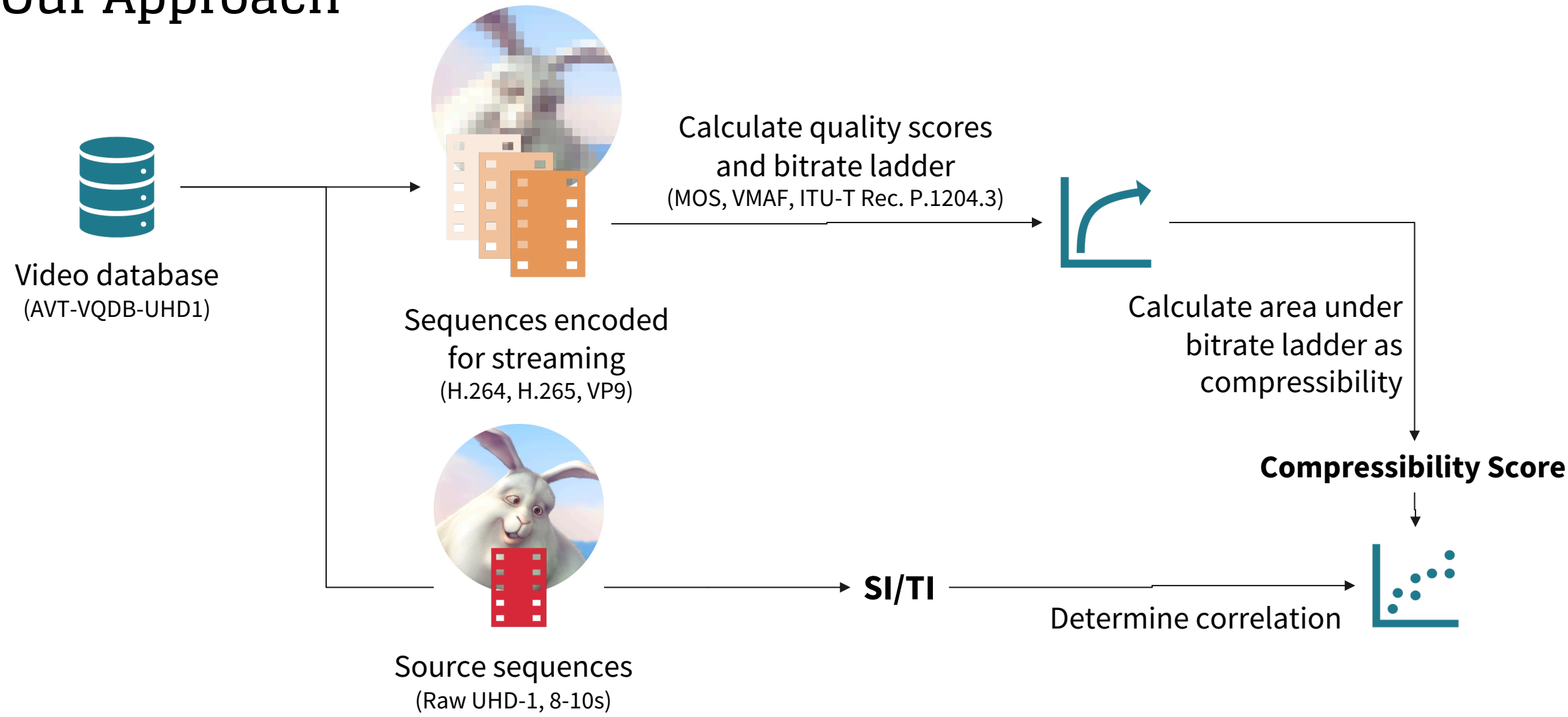
Videos with higher SI/TI have lower quality
when compressed under bitrate constraint

Compressibility of a source == achievable quality under bitrate constraints

High SI/TI lead to lower quality
and lower compressibility

Quality == subjective or “objective” MOS

Our Approach



Database and Features

AVT-VQDB-UHD1[1]

— Test 1: 180 PVSes

— Tests 2 & 3 with overlapping PVSes → mapped into one “virtual” test called “test 2+3” with 320 PVSes

— Test 4: includes FPS changes, therefore ignored

SI and TI scores

— Per frame, averages, minimum/maximum[2]

— Additionally: Criticality metric from Fenimore et al.[3] that uses SI/TI

$$C = log\left\{mean_{time}\left[SI\left(F_n\right) * TI\left(F_n\right)\right]\right\}$$

Video quality scores

— Subjective MOS from ITU-T Rec. P.910-compliant lab tests

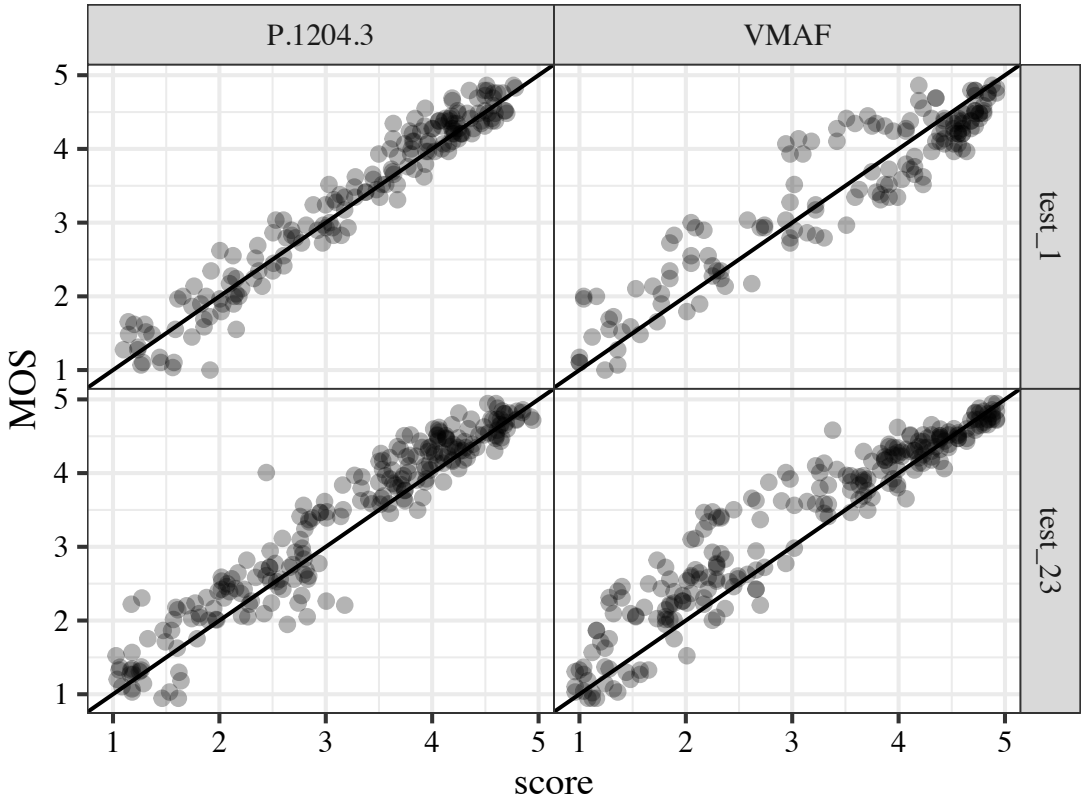
— VMAF (4K model, v0.6.1); full-reference model

— ITU-T Rec. P.1204.3; bitstream-based model[4]

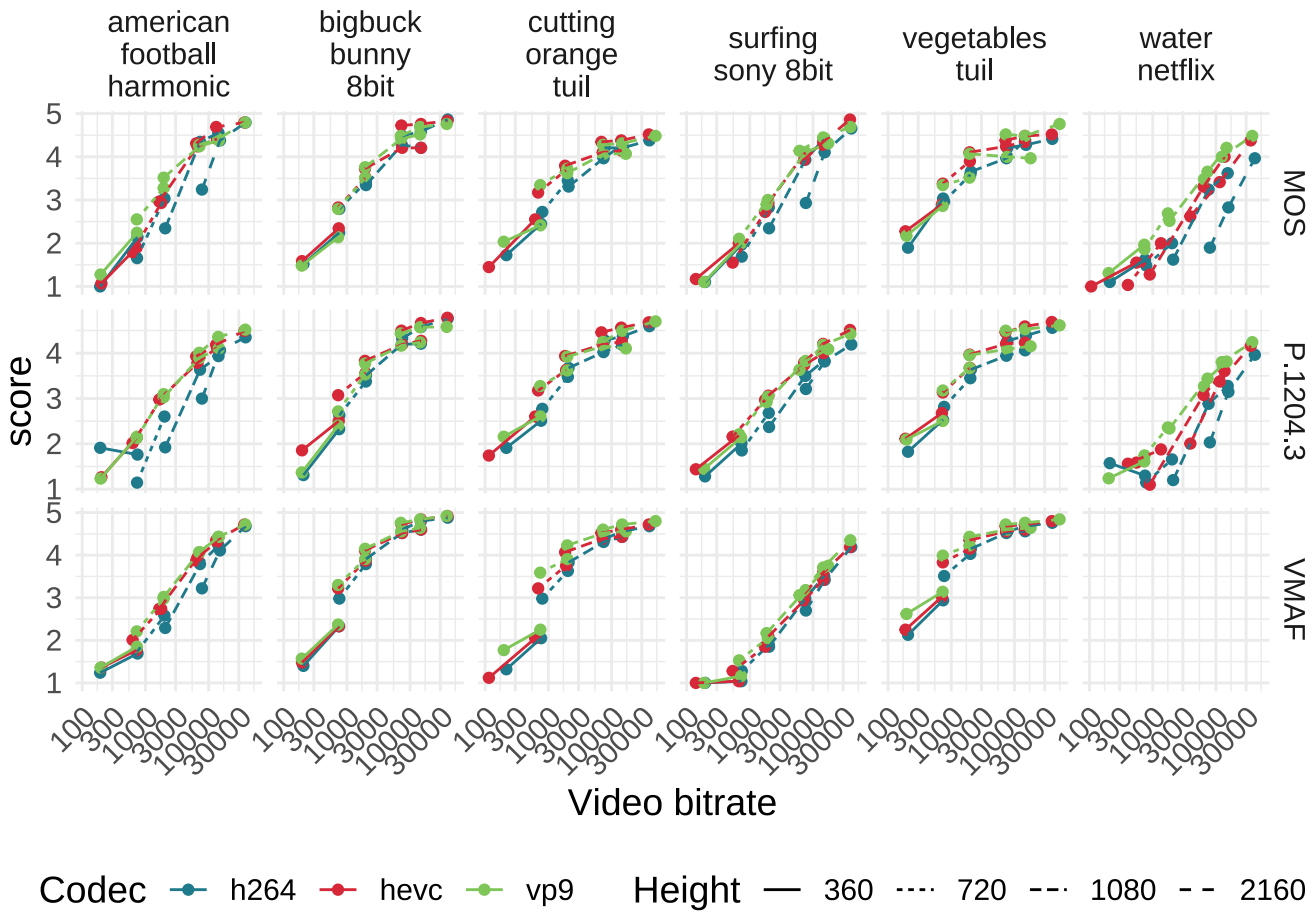
[1] https://github.com/Telecommunication-Telemedia-Assessment/AVT-VQDB-UHD-1
[2] calculated via https://github.com/Telecommunication-Telemedia-Assessment/siti-tools
[3] see Fenimore et al. (1998), Perceptual Effects of Noise in Digital Video Compression
[4] calculated via https://github.com/Telecommunication-Telemedia-Assessment/bitstream_mode3_p1204_3

Data Preparations

- Consistency checks:
  - "Dancers_8s" contained possibly erroneous PVSes already at encoding stage, and was completely removed
  - All VMAF scores for "water_netflix" sequence were erroneous and removed (possible frame offset)
  - P.1204.3 scores for four PVSes were removed due to being extreme outliers (possible model bug)
- How well do VMAF and P.1204.3 work?
  - Very well — see plot on the right
  - Pearson correlation between metric and MOS:
    P.1204.3: 0.96
    VMAF: 0.94

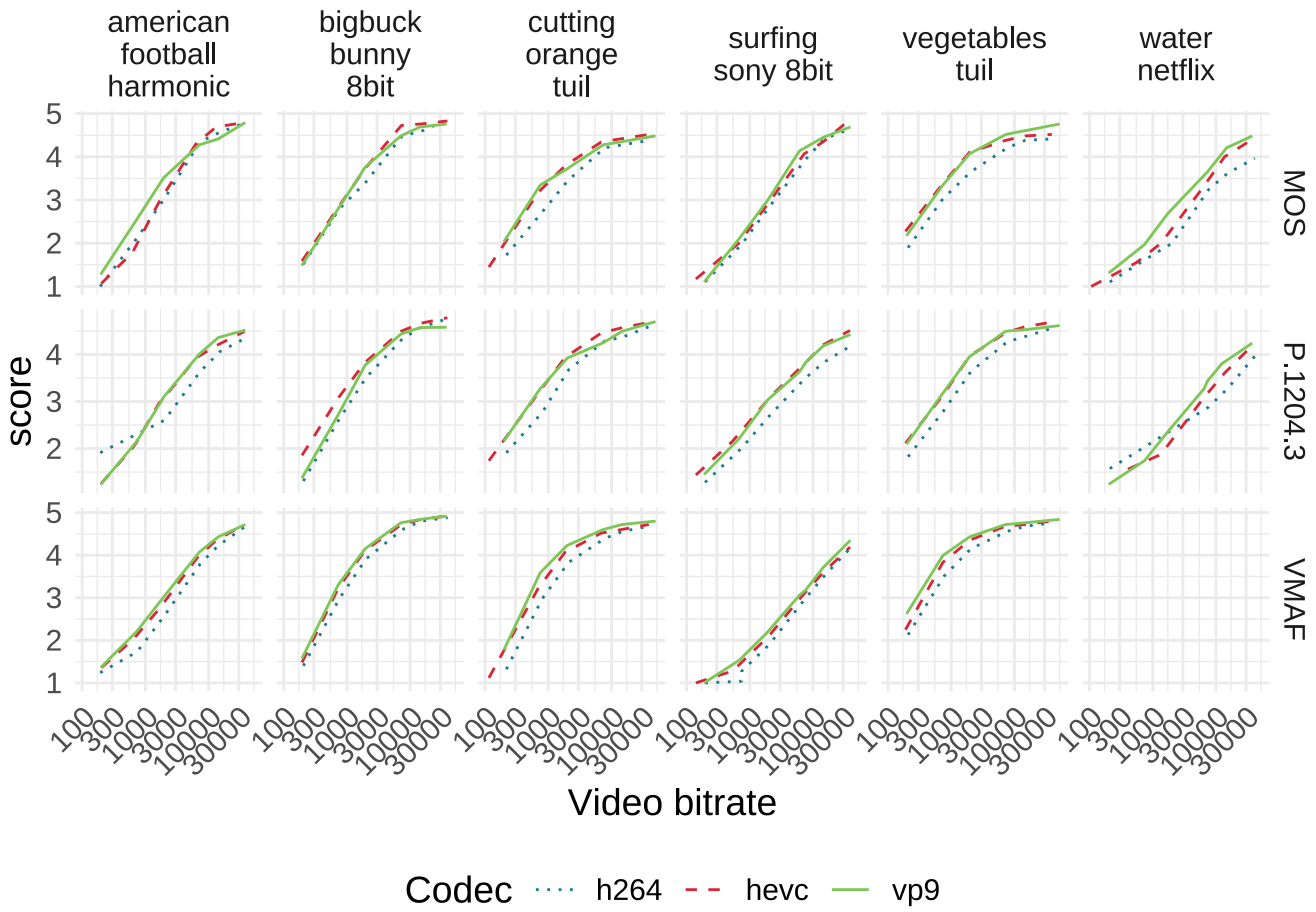


Determining Quality/Compressibility



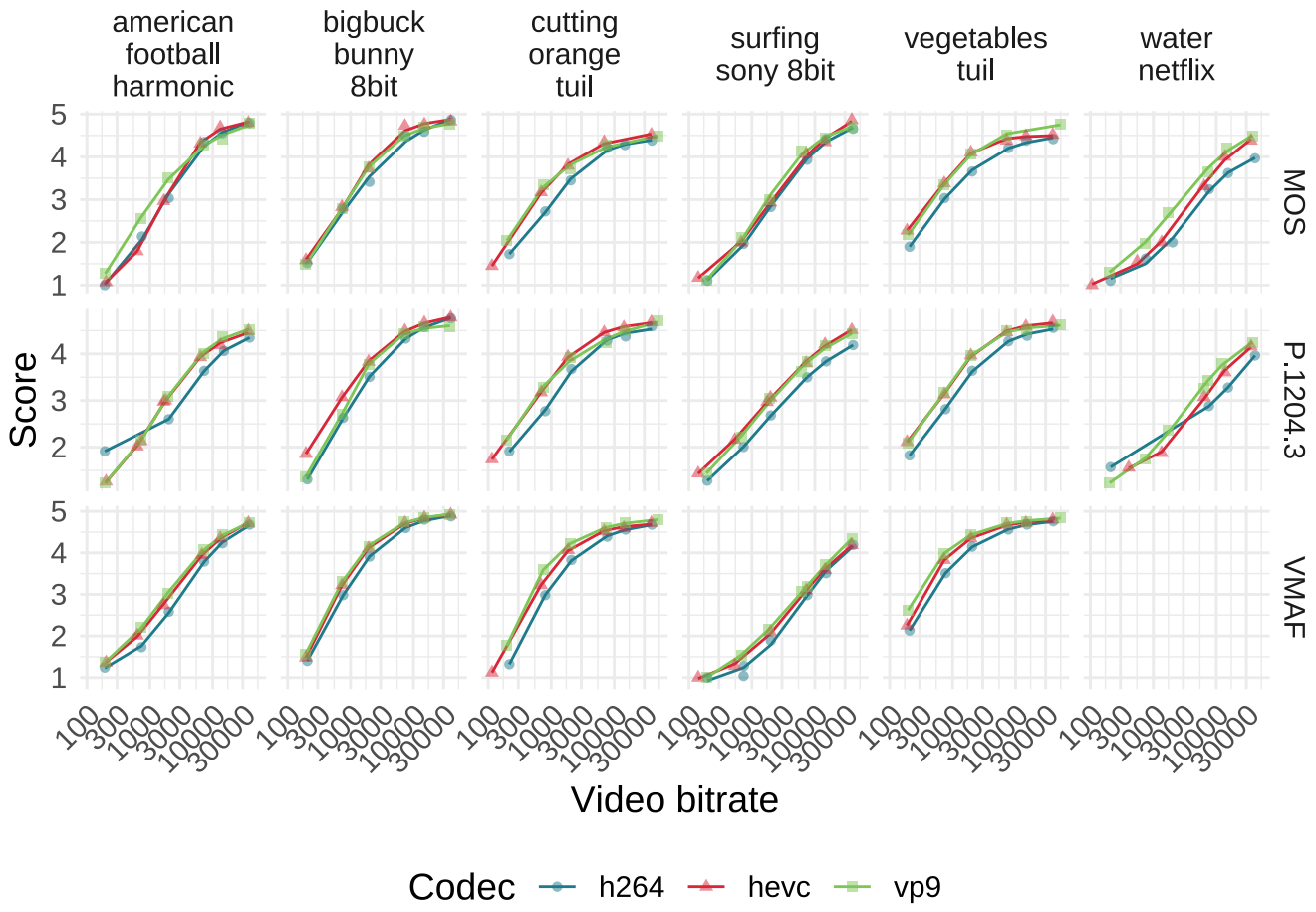
1. Create bitrate ladder for each SRC, i.e. MOS against bitrate

Determining Quality/Compressibility



1. Create bitrate ladder for each SRC, i.e. MOS against bitrate
2. Construct convex hull of ideal points

Determining Quality/Compressibility

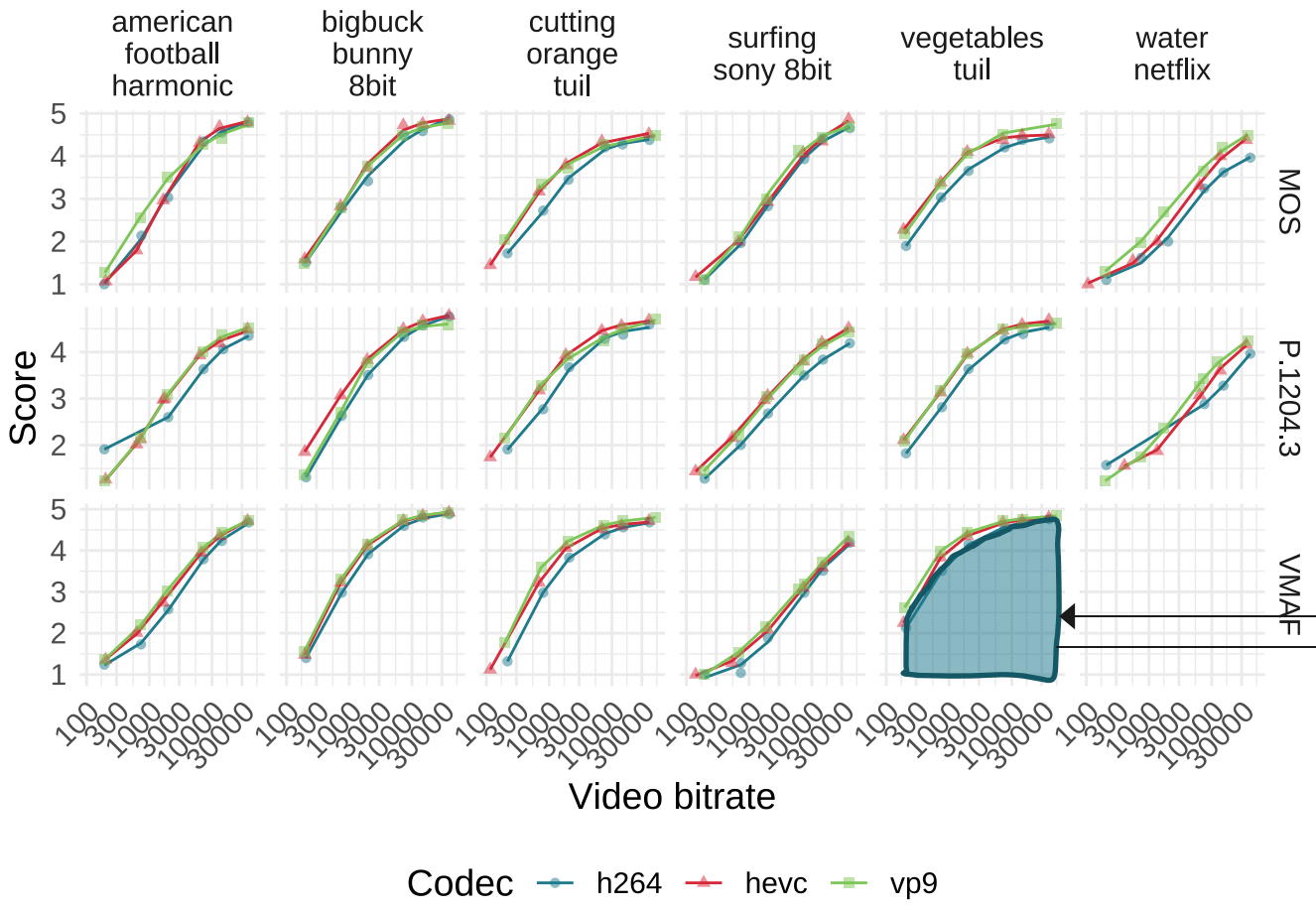


1. Create bitrate ladder for each SRC, i.e. MOS against bitrate
2. Construct convex hull of ideal points
3. Fit sigmoid function[1] against hull

$$S = a + \frac{b - a}{1 + exp\left(-c * (log(R) - d)\right)}$$

[1] based on: Hanhart et al. (2014), Calculation of average coding efficiency based on subjective quality scores

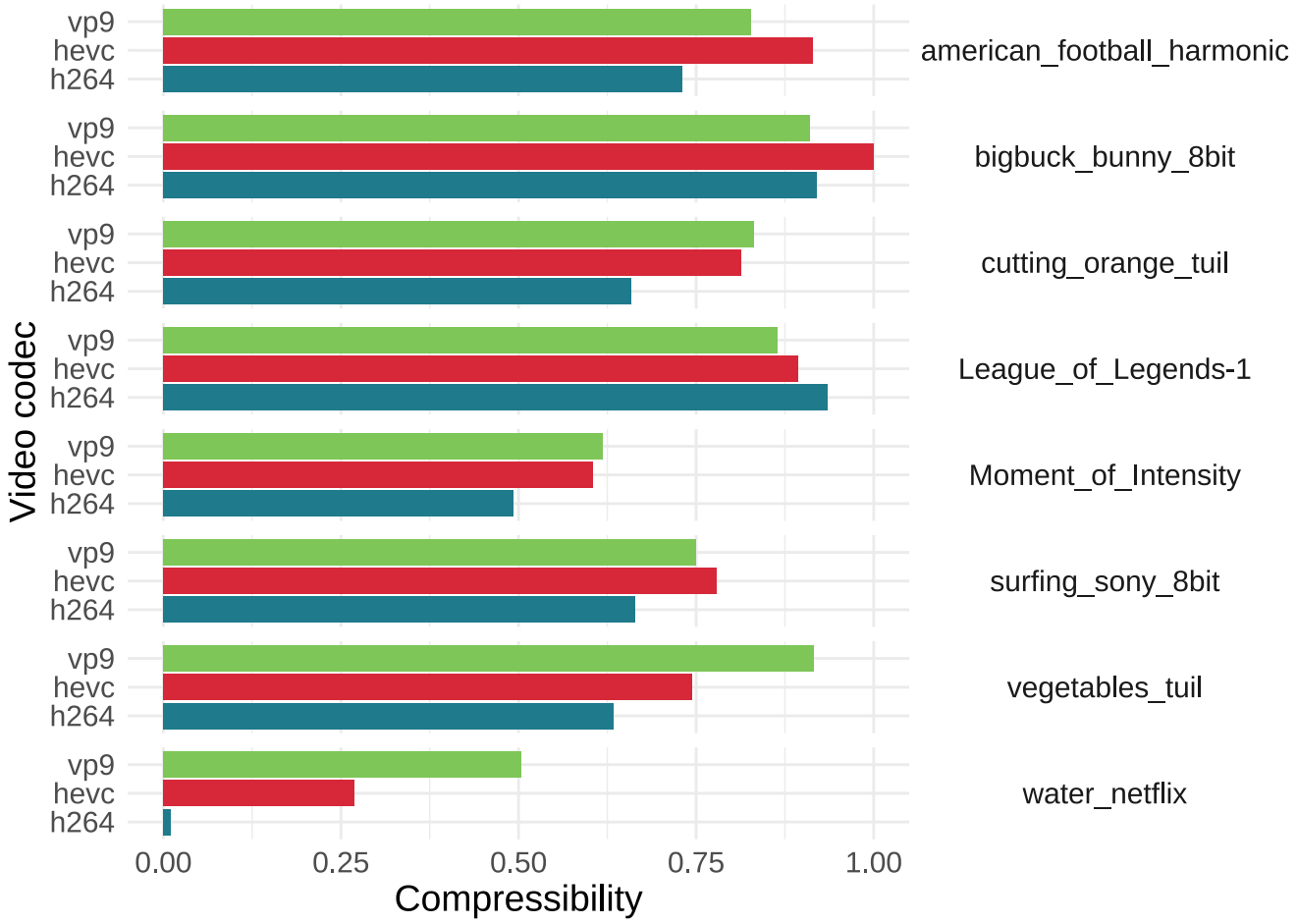
Determining Quality/Compressibility



1. Create bitrate ladder for each SRC, i.e. MOS against bitrate
2. Construct convex hull of ideal points
3. Fit sigmoid function against hull
4. Calculate are under each curve as quality/compressibility for that SRC/codec

Compressibility
1 number for each SRC/codec combination

Compressibility



→ Each SRC and codec have different compressibility scores

Note:

— Scores are normalized between 0 and 1 for this analysis (to be refined)

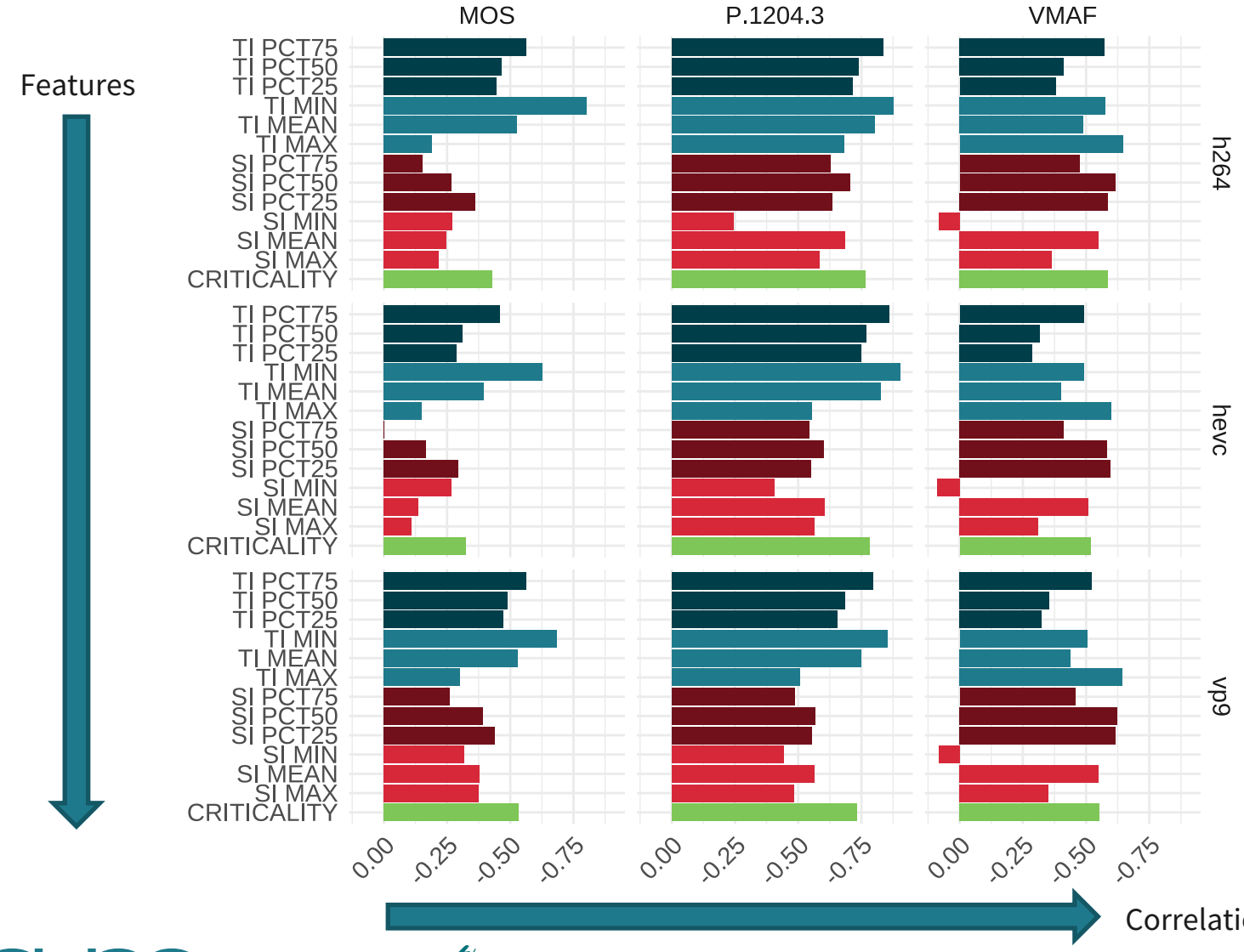
— Scores are based on MOS here

Examples:

— BBB is the easiest to compress, although used very often in tests

— Netflix Water sequence is the hardest

Correlations



- **TI features** have higher correlation with compressibility than **SI features**
- **Minimum TI** seems like a better correlated indicator than maximum TI
- **Mean SI** correlates better than min/max
- **Criticality metric** from Fenimore et al. has good correlation with compressibility

Summary

- Method to determine quality/compressibility of a given SRC and codec
  - Construct convex hull
  - Determine area under the curve
- Results:
  - SI/TI correlate with quality and compressibility
  - Minimum TI is a useful indicator too (for short sequences)
- Further research:
  - Are there combinations of SI/TI that can be more useful?
  - Determine compressibility as a single number on universal scale

Thank you!