Exploring the Relationship Between VQA Metric Accuracy and Bitrate Saving

Hojat Yeganeh, Ph.D, SMIEEE - Principal Video Architect and Research Lead

June 2023

Established success criteria for designing or selecting an objective VQA metric

When designing an objective IQA/VQA metric, the goal is to have maximum possible agreement between the metric and the subjective data (MOS) using the following criteria:

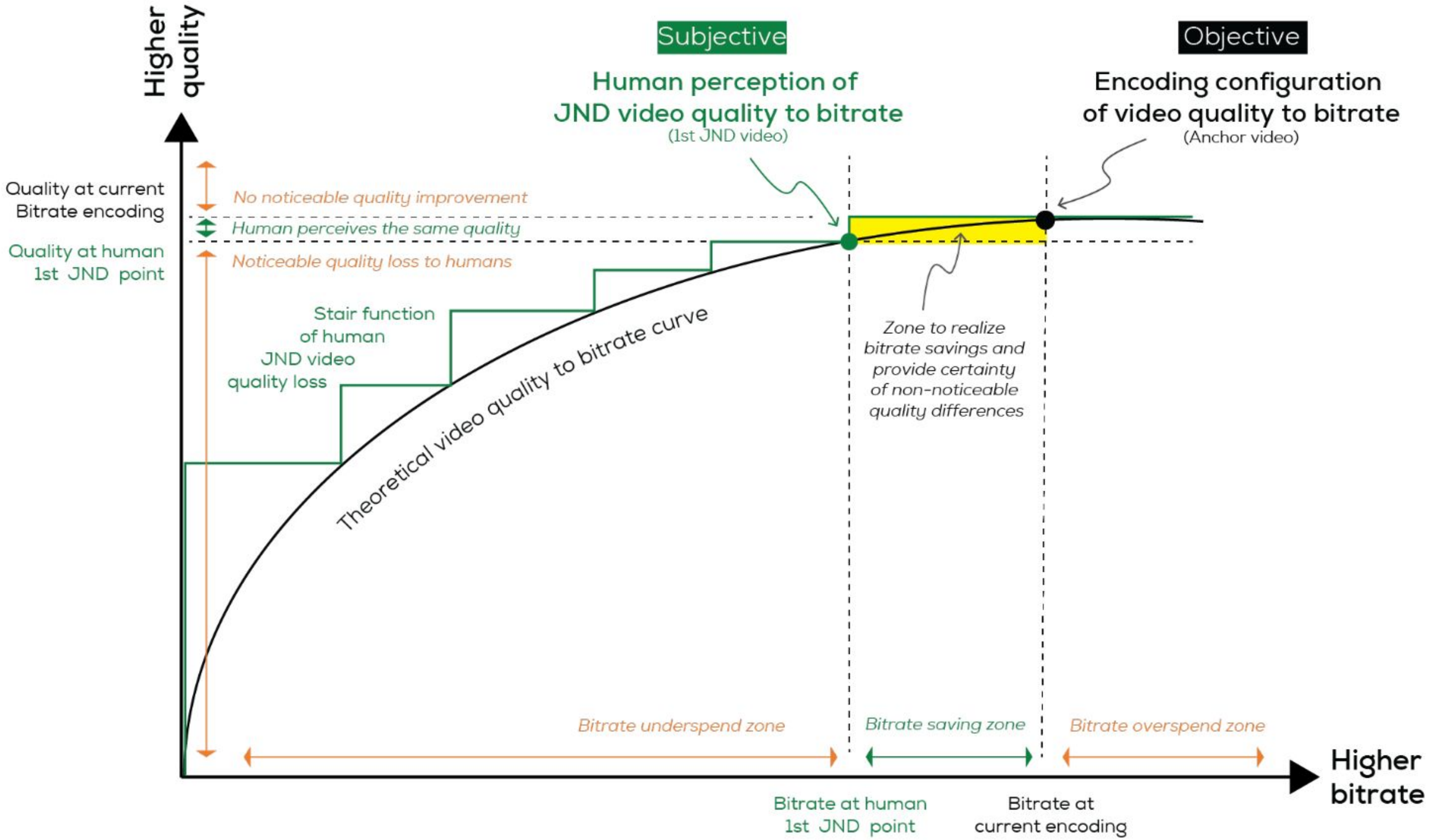
- Linear correlation (PLCC)
- Rank order correlation (SRCC, KRCC)
- Mean Absolute Error (MAE)
- Root Mean Square Error (RMSE)

Questions

- What is the purpose of selecting an objective VQA metric?
- Is a high correlation with subjective results a **necessary** or **sufficient** condition for a given use case?
- Specifically, for the **encoding optimization** use case, is having a high correlation with subjective results a necessary or sufficient condition?

Big decision

- In the video industry, it is a common practice to **optimize bandwidth usage** by decreasing the quality of videos, resulting in lower bit consumption. There's always a cost to compression.
- Achieving this goal is facilitated by employing any objective video quality metrics. VQA metrics are used as a threshold when reducing bitrate.
- This scenario remains applicable regardless of whether the optimization is performed per-segment, per-shot, per-scene, or per-asset.
- **The crucial question** lies in determining **the acceptable extent of quality reduction** and how an objective VQA model can predict such level of quality degradation.


Higher quality
Subjective
Human perception of
JND video quality to bitrate
(1st JND video)
Objective
Encoding configuration
of video quality to bitrate
(Anchor video)
Quality at current
Bitrate encoding
No noticeable quality improvement
Human perceives the same quality
Quality at human
1st JND point
Noticeable quality loss to humans
Stair function
of human
JND video
quality loss
Theoretical video quality to bitrate curve
Zone to realize
bitrate savings and
provide certainty
of non-noticeable
quality differences
Bitrate underspend zone
Bitrate saving zone
Bitrate overspend zone
Higher
bitrate
Bitrate at human
1st JND point
Bitrate at
current encoding

Identifying JND for a wide range of videos

VideoSet: A Large-Scale Compressed Video Quality DataSet Based on JND Measurement

VideoSet: A Large-Scale Compressed Video Quality Dataset Based on JND Measurement

Haiqiang Wang[1], Ioannis Katsavounidis[2], Jiantong Zhou[3], Jeonghoon Park[4], Shawmin Lei[5], Xin Zhou[6], Man-On Pun[7], Xin Jin[8], Ronggang Wang[9], Xu Wang[10], Yun Zhang[11], Jiwu Huang[10], Sam Kwong[12], and C.-C. Jay Kuo[1]

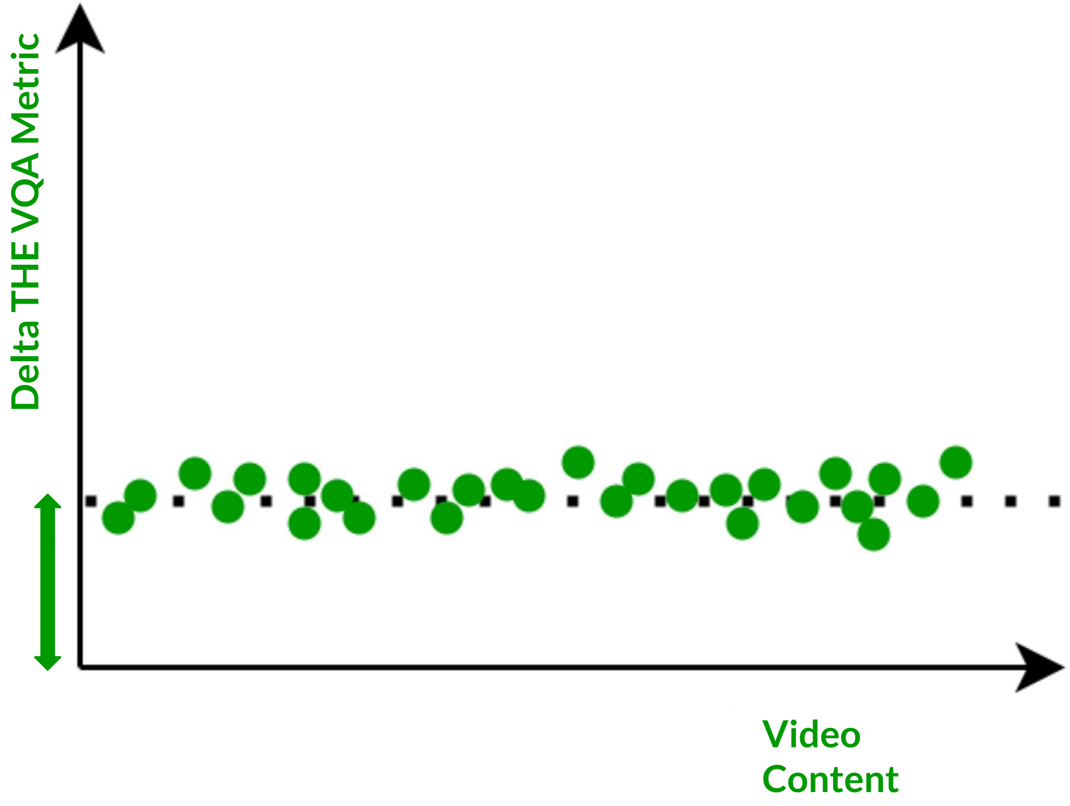
[1] University of Southern California, Los Angeles, California, USA
[2] Netflix, Los Gatos, California, USA
[3] Huawei Technologies, Shenzhen, China
[4] Samsung, DMC R&D, Seoul, Korea
[5] Mediatek, Hsinchu, Taiwan
[6] Northwestern Polytechnical University, Xi'an, China
[7] Chinese University of Hong Kong (SZ), Shenzhen, China
[8] Graduate School at Shenzhen, Tsinghua University, Guangdong, China
[9] Peking University Shenzhen Graduate School, Shenzhen, China
[10] Shenzhen University, Shenzhen, China
[11] Chinese Academy of Sciences, Shenzhen, China
[12] City University of Hong Kong, Hong Kong

Abstract. A new methodology to measure coded image/video quality using the just-noticeable-difference (JND) idea was proposed in [1]. Several small JND-based image/video quality datasets were released by the Media Communications Lab at the University of Southern California in [2,3]. In this work, we present an effort to build a large-scale JND-based coded video quality dataset. The dataset consists of 220 5-second sequences in four resolutions (i.e., 1920×1080, 1280×720, 960×540 and 640×360). For each of the 880 video clips, we encode it using the H.264 codec with $QP = 1, \cdots, 51$ and measure the first three JND points with 30+ subjects. The dataset is called the 'VideoSet', which is an acronym for 'Video Subject Evaluation Test (SET)'. This work describes the subjective test procedure, detection and removal of outlying measured data, and the properties of collected JND data. Finally, the significance and implications of the VideoSet to future video coding research and standardization efforts are pointed out. All source/coded video clips as well as measured JND data included in the VideoSet are available to the public in the IEEE DataPort [4].

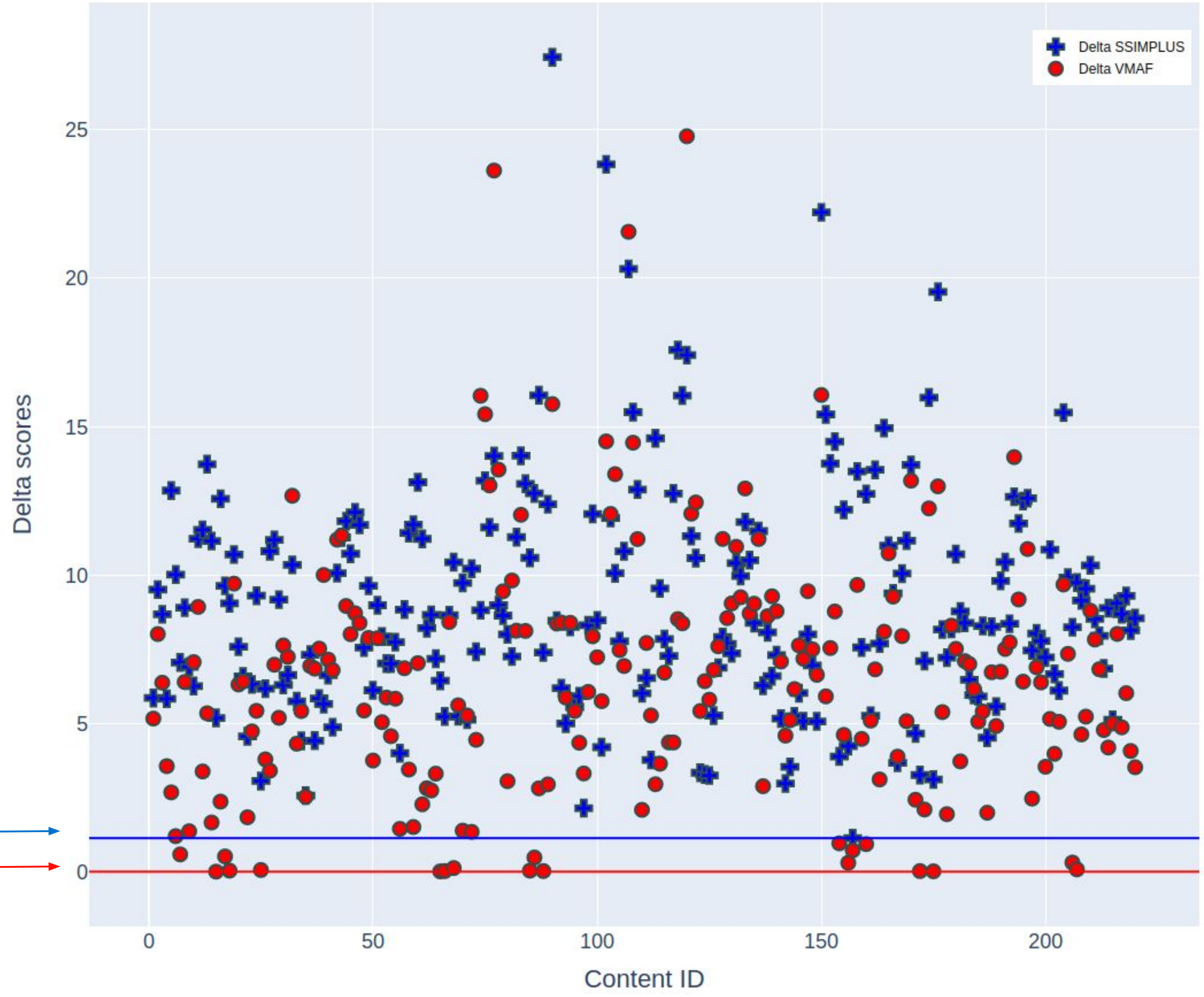
The expected response of a hypothetical perfect VQA metric to a JND dataset

- Delta THE VQA metric =
  metric (ref, anchor) - metric (ref, JND videos)

- JND videos are the videos that exhibit
  Just Noticeable Difference compared
  to the anchor videos

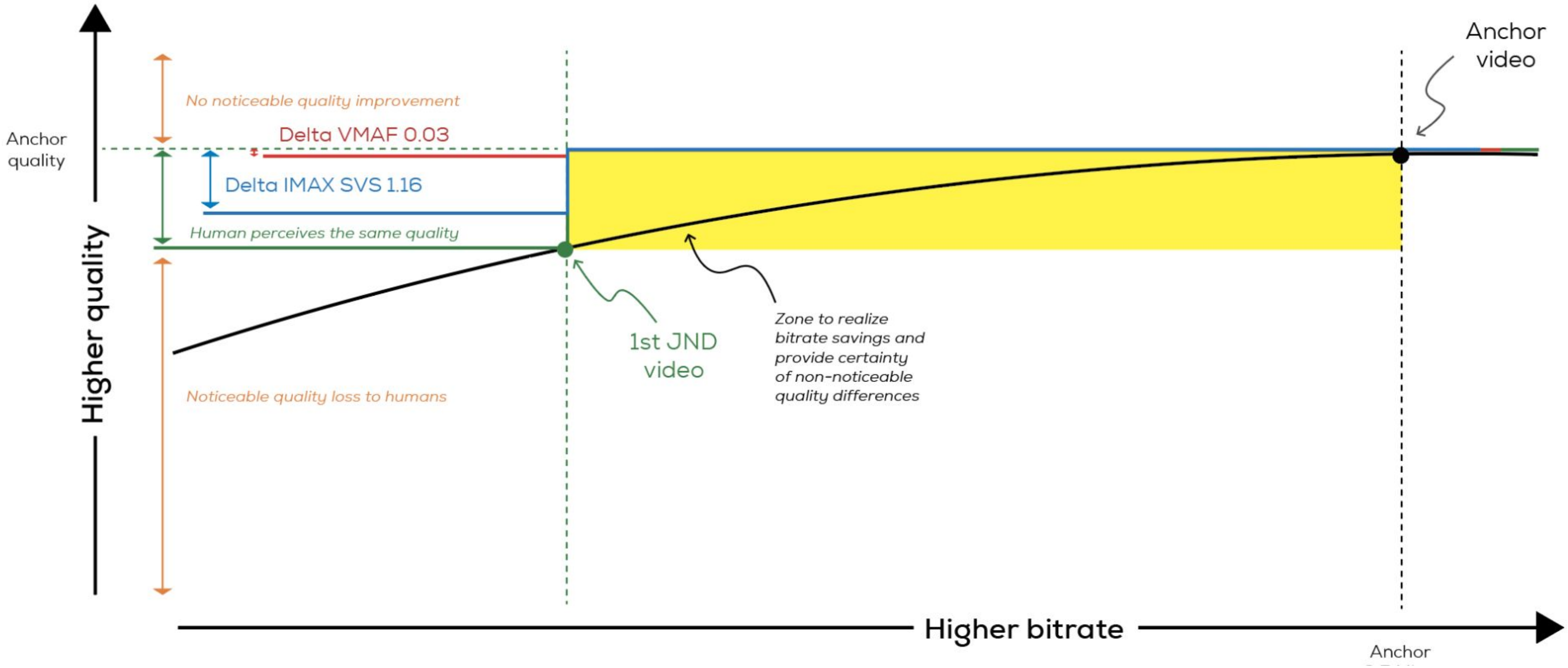


This figure illustrates the difference in **SSIMPLUS (IMAX SSIMPLUS Viewer Score)** and **VMAF** (developed by Netflix) scores between the anchors and the videos identified as the 1st Just Noticeable Difference (JND) in the VideoSet dataset.



Minimum Delta SSIMPLUS = 1.16

Minimum Delta VMAF = 0.03


Anchor video
No noticeable quality improvement
Delta VMAF 0.03
Anchor quality
Delta IMAX SVS 1.16
Human perceives the same quality
Higher quality
1st JND video
Zone to realize bitrate savings and provide certainty of non-noticeable quality differences
Noticeable quality loss to humans
Higher bitrate
Anchor

Significance of the minimum delta VQA with JND dataset

- The minimum delta VQA helps making a crucial decision when performing an encoding optimization
- Essentially, the minimum delta VQA metric using a JND dataset determines the level of quality degradation that can be tolerated without perceptible loss across a diverse range of video content
- Identifying this threshold is very important if the user cares about the **quality** of their content

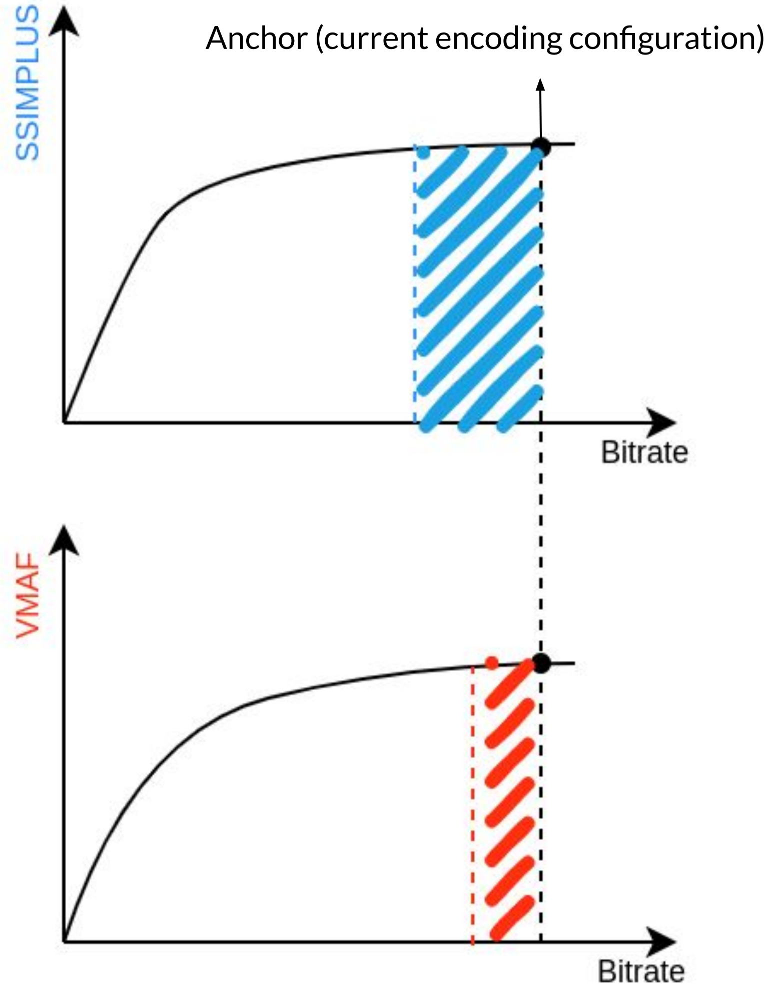
Metrics and savings

Assumptions:

- Both VMAF and SSIMPLUS satisfy necessary conditions, i.e.
  - a) Have high correlations with MOS
  - b) Have low MAE
  - c) Linear enough

Meaning: They give us more or less the same Q-D curves

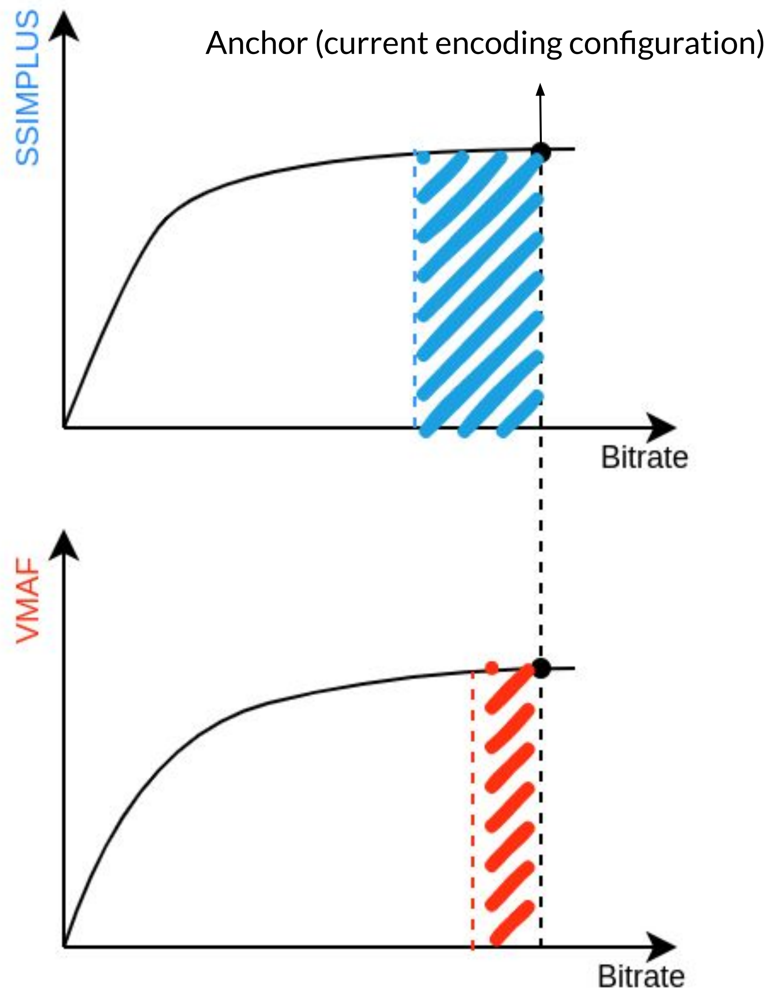
- Using JND data, there is lower room for quality reduction based on VMAF compared to SSIMPLUS



Metrics and savings

By specifically examining the first Just Noticeable Difference (JND) and gathering the corresponding bitrate values for those videos from the VideoSet dataset, it is possible to estimate the potential bandwidth savings using the two metrics with different anchor bitrate values.

Anchor bitrate (1080p)	Delta IMAX SVS 1.16	Delta VMAF 0.03
~8.5Mbps	Save 28.94%	Save 2.92%
~7.0Mbps	Save 25.80%	Save 1.91%
~6.5Mbps	Save 24.09%	Save 1.62%



Conclusion

- PLCC, SRCC, KRCC, MAE, RMSE, etc. serve as **necessary conditions** in selecting an objective VQA metric for **encoding optimization** use cases but **are not sufficient** on their own

- For making better encoding decisions, exploring the performance of an objective VQA model against a subjectively collected JND dataset is of utmost importance

- Understanding the minimum acceptable reduction in the objective VQA metric, which corresponds to just noticeable difference (JND) quality, becomes imperative when preserving the level of perceptual quality is a priority

- A higher value for the minimum acceptable reduction based on an objective VQA metric can potentially result in greater bitrate savings, as it provides more flexibility to optimize encoding parameters and achieve significant reductions in bitrate

Thank you

Hojat Yeganeh

hojat.yeganeh@ssimwave.com