

Measuring Video Quality with VMAF: Why You Should Care

Christos Bampis
Encoding Technologies, Netflix

AOMedia Research Symposium
San Francisco, October 15, 2019

NETFLIX

Overview

- history and introduction to VMAF
- adoption
- challenges
- why is VMAF becoming more useful?

Need a better perceptual metric



PSNR 29.1 dB

19



PSNR 29.3 dB

69

Humans

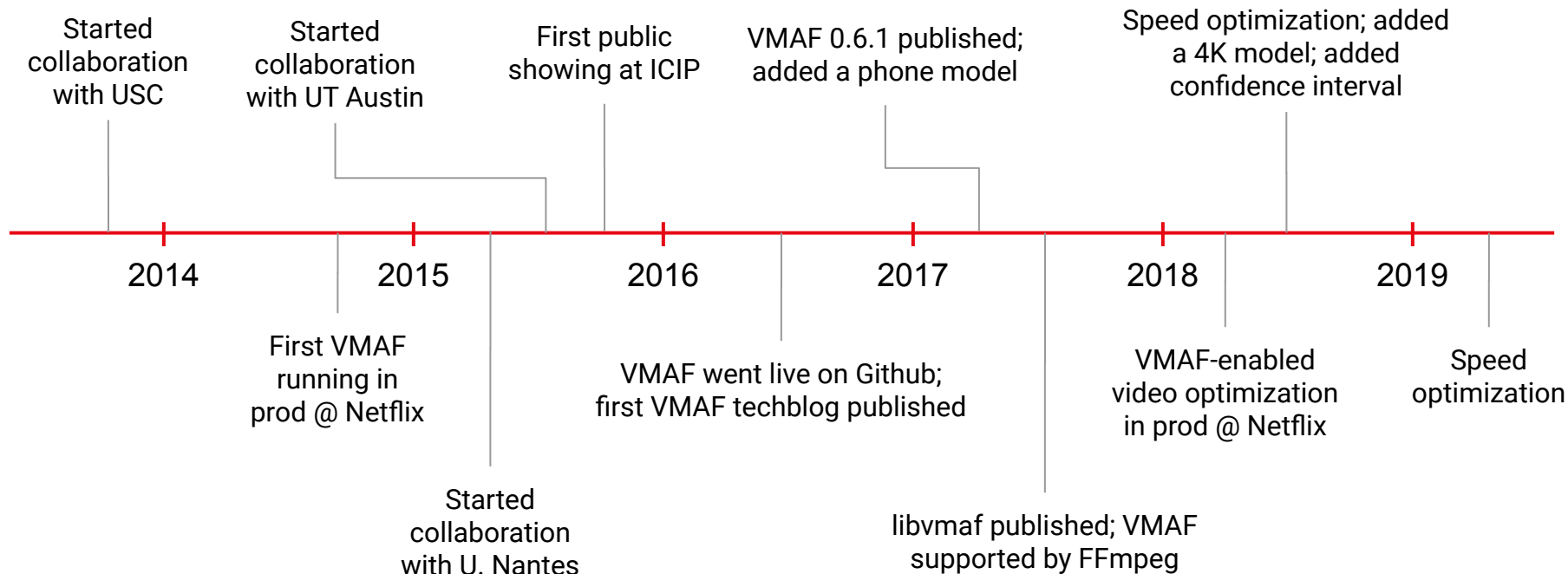
VMAF

- accurately measures **human perception** of video quality
- consistent across content
- works well for picture artifacts relevant to adaptive streaming
 - compression artifacts
 - scaling artifacts
- open-source!

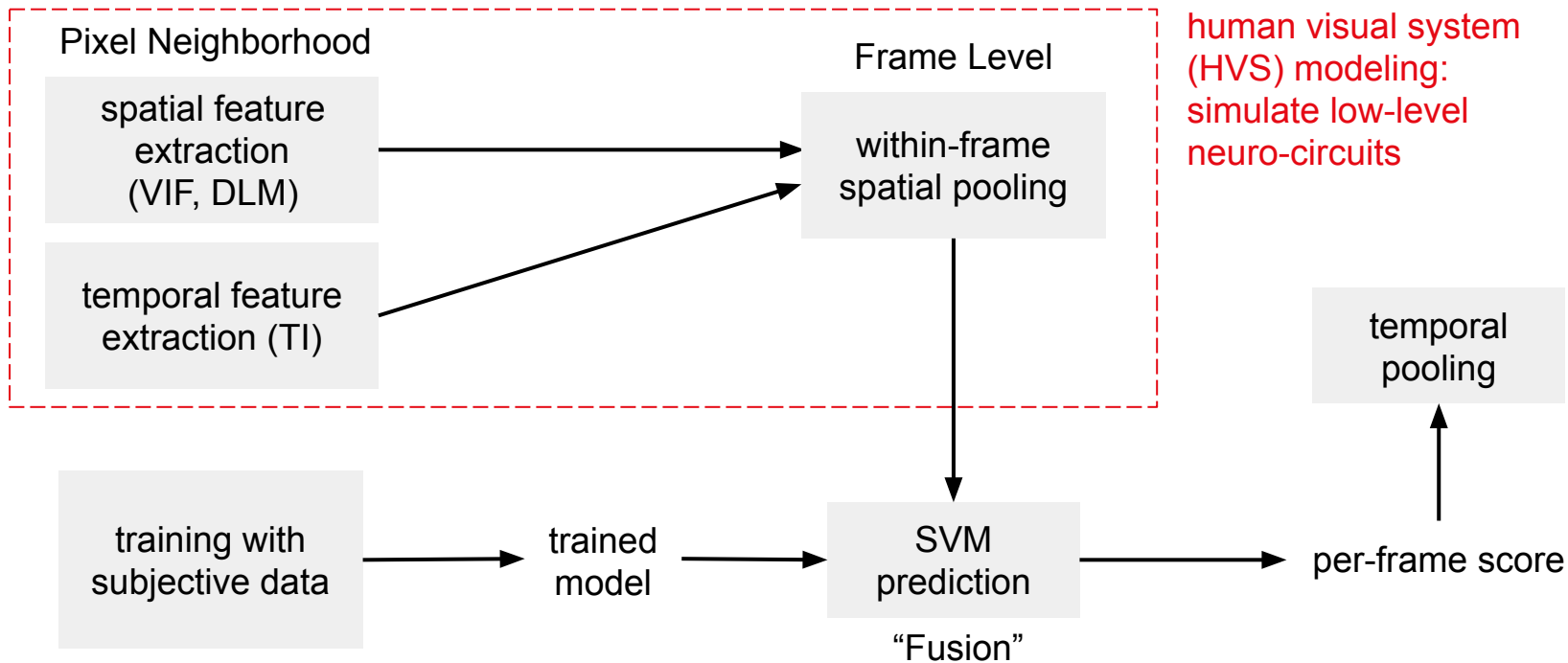


VMAF: **V**ideo **M**ultimethod **A**ssessment **F**usion

The VMAF chronicle

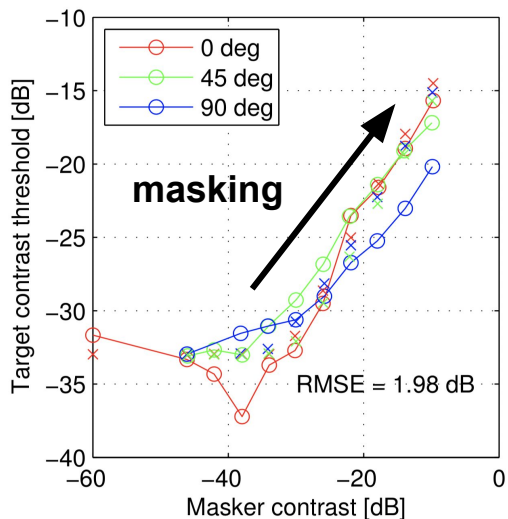


VMAF framework



HVS modeling: contrast masking

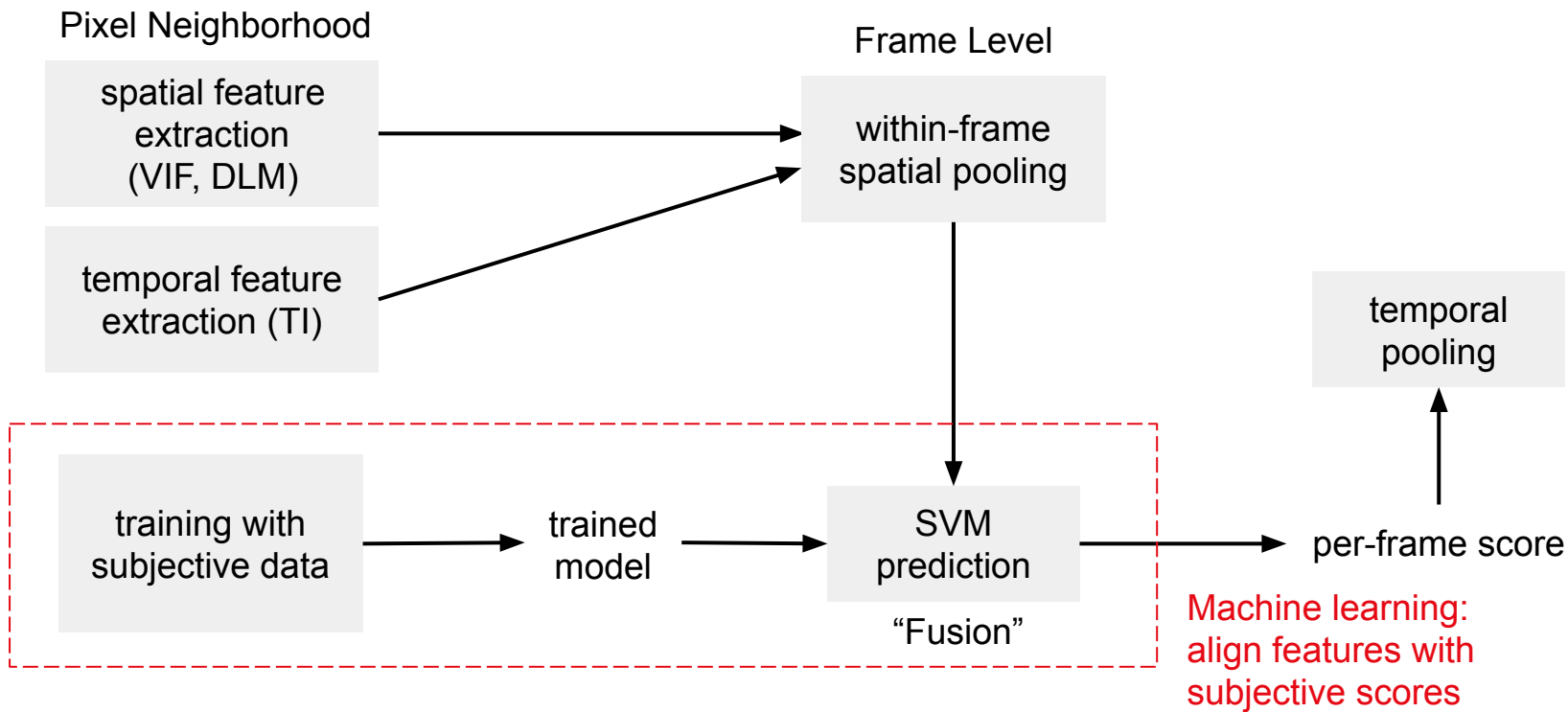
- One signal (e.g. compression artifacts) becomes more difficult to be detected by human eye when it is superimposed on another masker signal (e.g. the pristine source) of similar spatial frequency and orientation



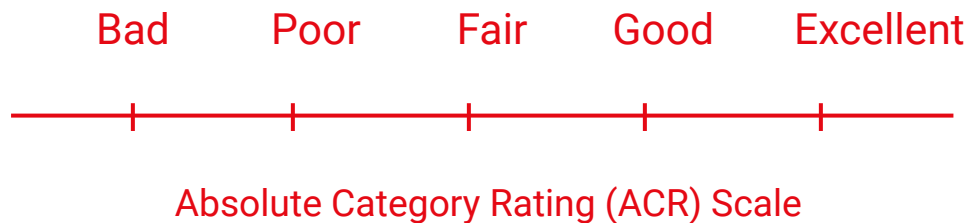
[Source: HDR-VDP2, Mantiuk et al. 2011]



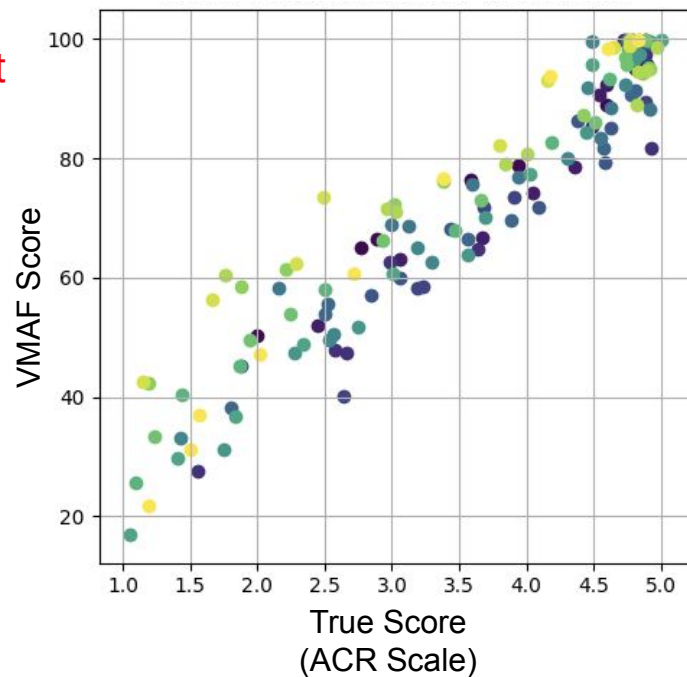
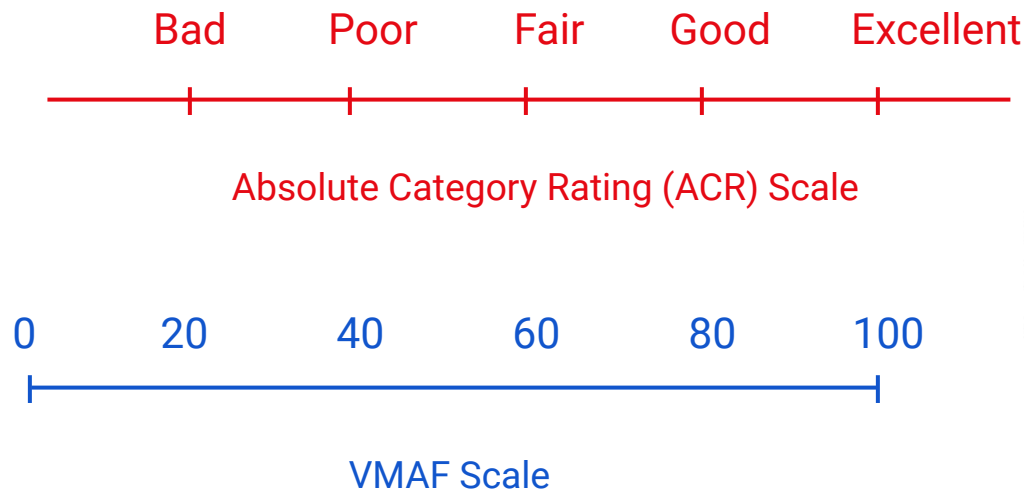
VMAF framework



Lab test: collect subjective scores



Map ACR scale to VMAF scale



Demo time!

VMAF adoption examples

- industry
- research community

Integration in 3rd-party tools

 FFmpeg

About

News

Download

Documentation

Community

Mailing Lists

IRC

Forums

Bug Reports

Wiki

Develop

Source

10.107 libvmaf

Obtain the VMAF (Video Multi-Frame Analysis) score for your videos.

The obtained VMAF score is printed to the console.

It requires Netflix's vmaf library. It can be enabled using the `--enable-libvmaf` option. If no model path is specified, the default model path is used.

The filter has following options:

model_path
Set the model path with the `-model_path` option. Example: `-model_path "vmaf_v0.6.1.pkl"`




log_path
Set the log path with the `-log_path` option. Example: `-log_path "vmaf.log"`



VIDEO COMPRESSION GURU

Video Quality Measurement Tool

Video Quality measurement tool designed to compare quality of encoded streams based on objective metrics, such as PSNR, APSNR, SSIM, DELTA, MSE, MSAD, VQM, NQI, VMAF and VMAF phone.

[Contact Us](#) [Try for Free](#)   

MSU Video Quality Measurement Tool

[MSU Graphics & Media Lab \(Video Group\)](#)

Implementation: Alexey Moskvina, Oleg Petrov, Sergey Putilin, Sergey Grishin, Arseniy Georgiyev

Projects, ideas: Dr. Dmitry

▼ Why VQMT?

- [Basic Information](#)
- [Why update?](#)
- [PRO Version \(with command-line interface\) benefits](#) [Full]
- [Key Contributors](#)

▼ Support

- [Setting & Screenshots](#) (advanced files settings, visualization settings)
- [Metrics Info](#) (VMAF, NIQE, PSNR, MSE, MSAD, SSIM, VQM, MSU ETC)
- [FAQ](#) (Frequently Asked Questions)
- [Plugins & Plugin SDK](#)
- [MSU VQMT program performance](#)
- [Subjective quality metrics comparison](#)

ZOND 265 — your indispensable tool

- ✓ Visualized bits distribution statistics to facilitate encoding of CU, PU, and TU data
- ✓ Quality estimation by Netflix VMAF, PSNR and MSSIM metrics
- ✓ HEVC stream validation conformance according to the specification in GUI
- ✓ Fast translucent overlays of encoded file image (CU, PU, TU, QR, Bits, Refldx, PNSR) for Ultra HD files.

 VideoClarity

CONTACT US

Tools for Video Analysis

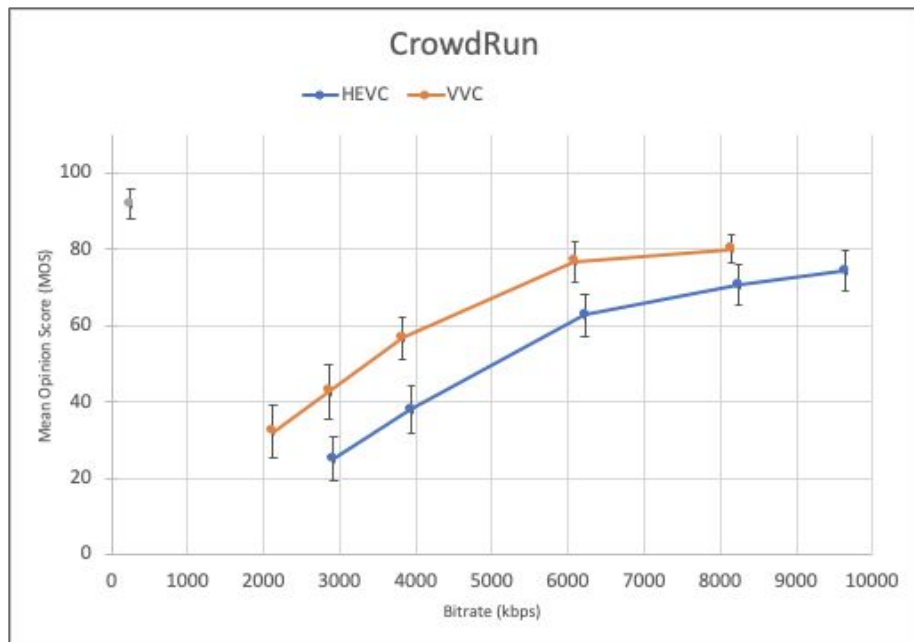
HOME **PRODUCTS** SOLUTIONS INFORMATION SUPPORT OFFICES

Automated Testing Solutions

ClearView – Video Quality Measurement and Analysis Systems

- Resolution and frame rate independent picture quality analyzers with multiple models
- Versatile input/output with 12G-SDI, 10G IP, HDMI, IP stream decode/capture
- Comprehensive media ingest with full featured file decoder application
- Quantitative picture quality metrics VMAF, NIQE, MS-SSIM/DMOS, PSNR, and Sarnoff JND (option)
- Test for HDR color just noticeable differences with ΔE_{ITP}
- Test for source sequence complexity using Temporal and Spatial metrics
- Performance testing for audio using aFreq metric for audio quality with A/V offset (lip-sync) measurement and aPeak metric for standardized loudness testing
- Automatic temporal and spatial alignment, comparison playback of uncompressed video via HDSI or HDMI outputs with interactive side-by-side views up to 8K and 4K @ 60 Hz
- Picture zoom up to 16X with pan and scroll provides unmatched inspection of compression artifacts during any play mode for a detailed view of one or a comparison of two videos
- Desktop window of comparison play modes provides for machine room or data center use without video output connectivity

VMAF in codec comparisons



Resolution	BD-rate (PSNR)	BD-rate (VMAF)	BD-rate (MOS)
HD	-31.24%	-35.18%	-36%
UHD	-34.42%	-40.44%	-40%

[Source: JVET-O0451 Subjective Comparison of VVC and HEVC, JVET 15th meeting: Gothenburg, SE, 3–12 July 2019]

VMAF in research papers

VMAF reproducibility: Validating a perceptual practical video quality metric

R Rassool - 2017 IEEE International Symposium on ..., 2017 - ieeexplore.ieee.org

Measuring video quality with standard metrics ensures that operators can deliver to consumers the desired quality of experience (QoE) at an optimal cost. Such metrics also

Video Multimethod Assessment Fusion (VMAF) on 360VR contents

M Orduna, C Díaz, L Muñoz, P Pérez, I Benito... - arXiv preprint arXiv ..., 2019 - arxiv.org

This paper describes the subjective experiments and subsequent analysis carried out to validate the application of one of the most robust and influential video quality metrics, Video

Practical Evaluation of VMAF Perceptual Video Quality for WebRTC Applications

B García, L López-Fernández, F Gortázar, M Gallego - Electronics, 2019 - mdpi.com

WebRTC is the umbrella term for several emergent technologies aimed to exchange real-time media in the Web. Like other media-related services, the perceived quality of WebRTC

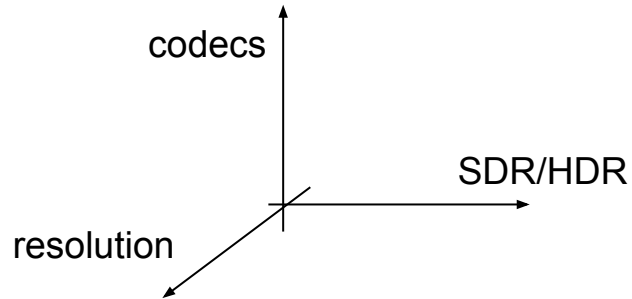
NR-GVQM: A no reference gaming video quality metric

S Zadtootaghaj, N Barman, S Schmidt... - ... on Multimedia (ISM), 2018 - ieeexplore.ieee.org

... NR-GVQM is designed by training a Support Vector Regression (SVR) with the Gaussian kernel using nine frame-level indexes such as naturalness and blockiness as input features and Video Multimethod Assessment Fusion (**VMAF**) scores as the ground truth ...

What are the challenges?

- design dimensionality
- dealing with noise



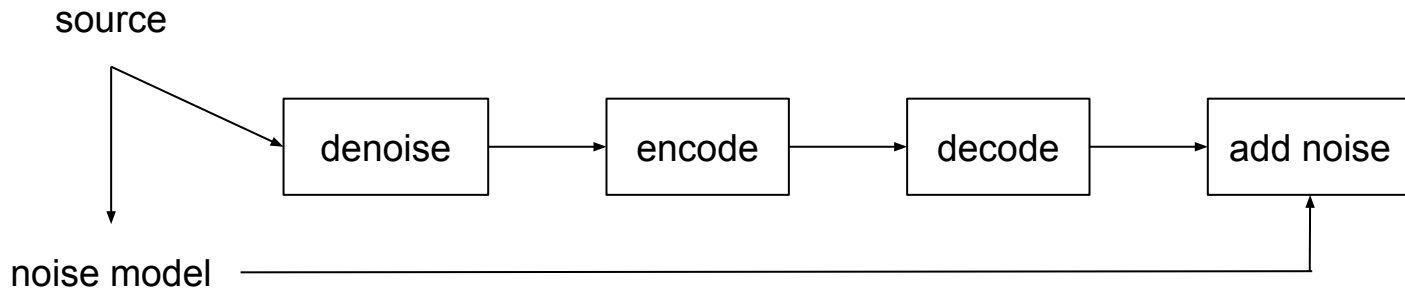
Design dimensionality

increased number of dimensions:

- different encoders: H.264/AVC, HEVC, VP9, AV1
- SDR vs. HDR, dark vs. bright scenes
- different viewing conditions (phone vs. TV, 1080 vs. 4K)
- key question: how to design a model that is extensible
and consistent?

Dealing with noise

- VMAF underpredicts under noisy source
- assess film-grain synthesis tools (e.g. AV1)



Why is VMAF becoming more useful?

- newer codecs (e.g. AV1) add more perceptual tools to their arsenal and PSNR is not enough to evaluate them
- open-source and well-adopted: problems are easier to find
- we are committed to further improving VMAF's accuracy and speed

Summary

- VMAF aims to fill the gap in perceptual video quality metrics
- adopted by industry and academia, but there is room for improvement
- becomes more relevant for new and future codecs (AV1, AV2), e.g., for codec comparison, encoding optimization

Questions?