The Demise of the Hybrid Codec?

Al Bovik

Who's Gonna Win?

Reigning champ: MPEG Hybrid

+ Enormous momentum
+ Big head start
+ Device compatibility
+ Continued "handcrafted" creativity
+ Compute gains driving exhaustive redundancy reduction

- Enormous (?) inertia
- Limits to "handcrafting"
- Lack of perceptual design
- Need lots of perceptual data

The Deep Learning Upstart

+ Big paradigm shift – nimble field
+ Performance already comparable
+ Can be done at scale (Google etc)
+ Millions of parameters
+ Compute gains driving optimization of millions of parameters
+ Perception-savvy crowd

- Unknown limits
- Device implementation HARD
- Need LOTS of perceptual data

How to (Try to) Save the Hybrid Codec

- More **perceptual design** (models and metrics - VMAF, MS-SSIM

- Machine learn **all adaptions**
  - mode decisions
  - QP levels
  - spatial resolution, etc

  while optimizing **perception**

- Test and report all comparisons using **perceptual metrics**, or when possible, **large human studies**

- **Lose PSNR**

- Simplify and **perceptually optimize** codec end-to-end, while maintaining **scalability edge.**

- Perhaps specify the **entire codec**

- **Perceptually optimize** codec design against **standard RD operating points**, i.e., bitrate ladder of QP/resolution

There is a **history** of **perceptual design** in **video coding**!

Adaptive Quantization of Picture Signals Using Spatial Masking

ARUN N. NETRAVALI AND BIRENDRA PRASADA, SENIOR MEMBER, IEEE

(1977)

How to (Try to) Beat the Hybrid Codec

- **Machine learn end-to-end**
- ~~**Motion estimation**~~ **not needed?**
- Send the decoder with the code hence

  ~~**Standard decoder**~~
- Continue **perceptual (+ data) designs!**
- **DATA**. Pioneer **large-scale human subjective** studies on **compression.** This will require **significant creativity.**
- **Continue using** and **reporting** all comparisons using **subjective metrics** (MS-SSIM, VMAF)
- Focus on **scalability** (especially decoder) – or **forget it!**

- Start thinking **beyond the codec**
  - Device deployments (ASICs, SOCs)
  - R-D usage scenarios / perceptual operating points
  - Design relative to bitrate ladders / other adaptation (HTTP-DASH, etc)
- **Take the lead** on
  - 360 video / VR / AR
  - 8K / 16K / 32K