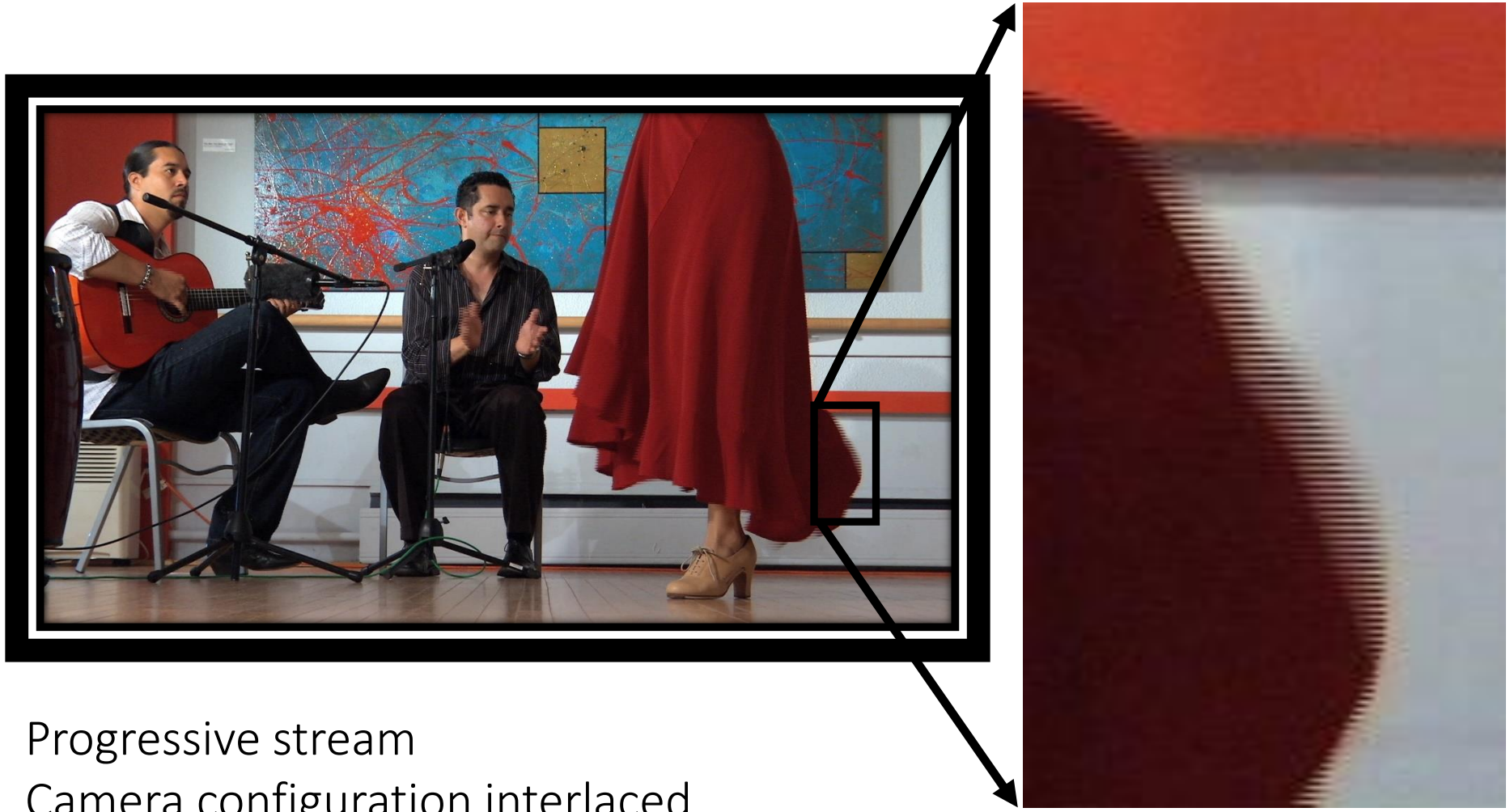
New paradigm to assess video quality in broadcast workflows

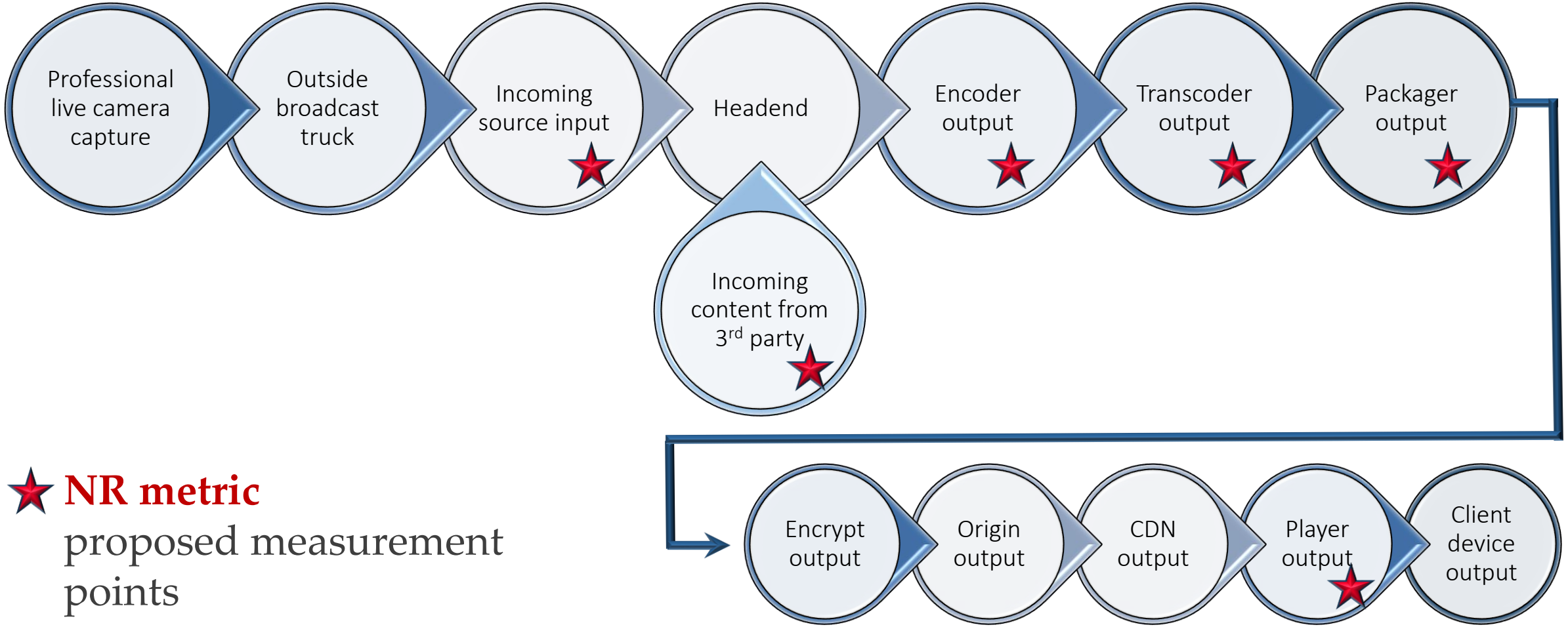
Margaret H. Pinson
mpinson@ntia.gov

ITS: The Nation's Spectrum and Communications Lab



Progressive stream
Camera configuration interlaced

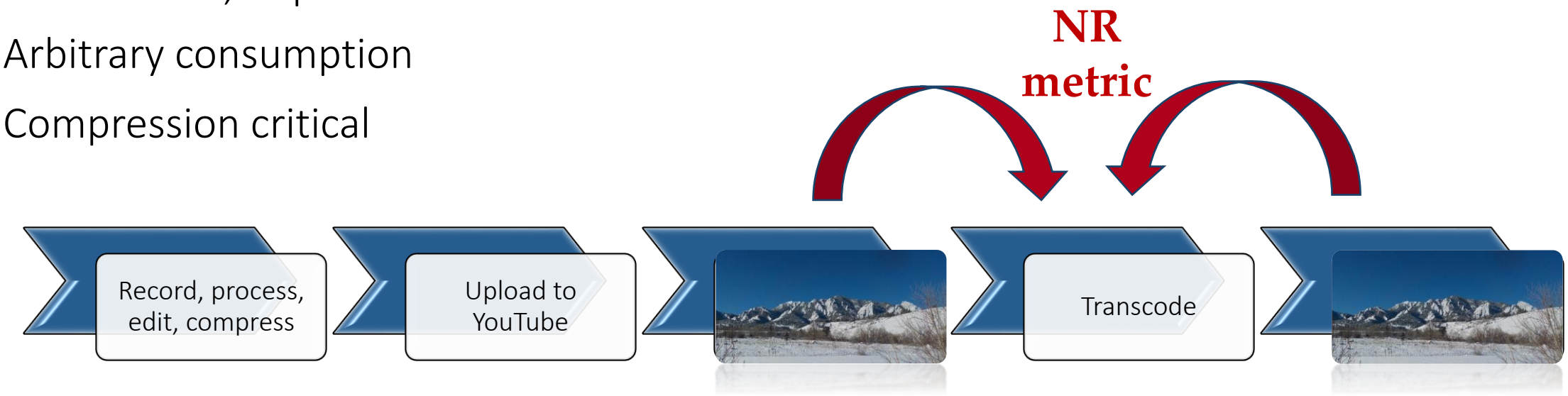
End-to-end Broadcast Workflow



★ **NR metric** proposed measurement points

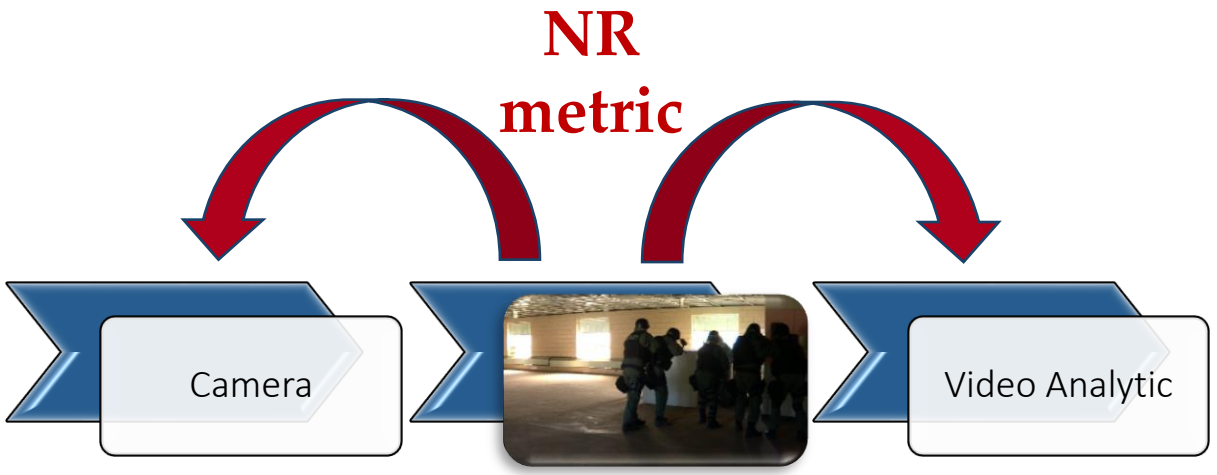
YouTube

- Millions of user generated contents each day
- Inconsistent, unpredictable media
- Arbitrary consumption
- Compression critical



Computer Vision

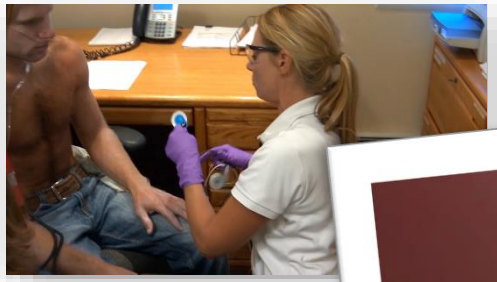
- Real-time camera feedback loop
  - Focus
  - Zoom
  - Exposure
  - Bit-rate
- Predict success / failure
- Choose between several algorithms

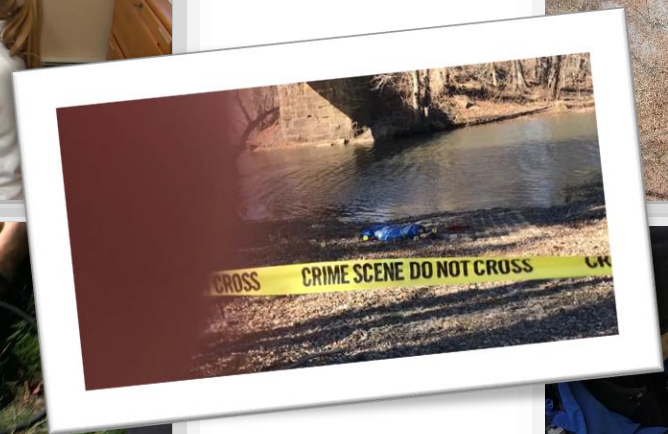


First Responders

- Transient event & difficult environment
- Task specific differences
- Different quality requirements

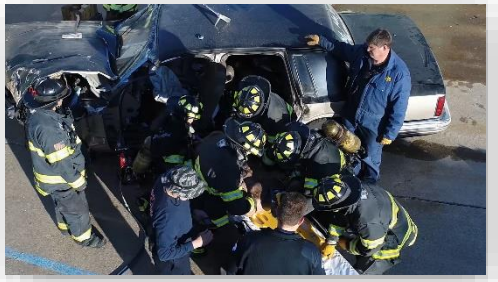




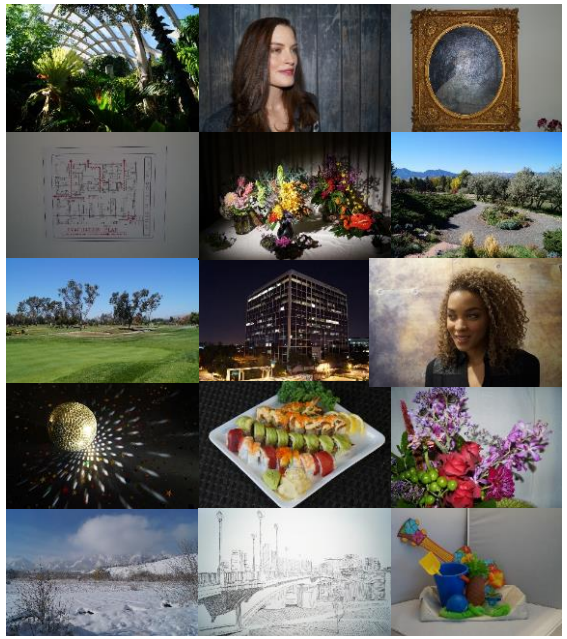









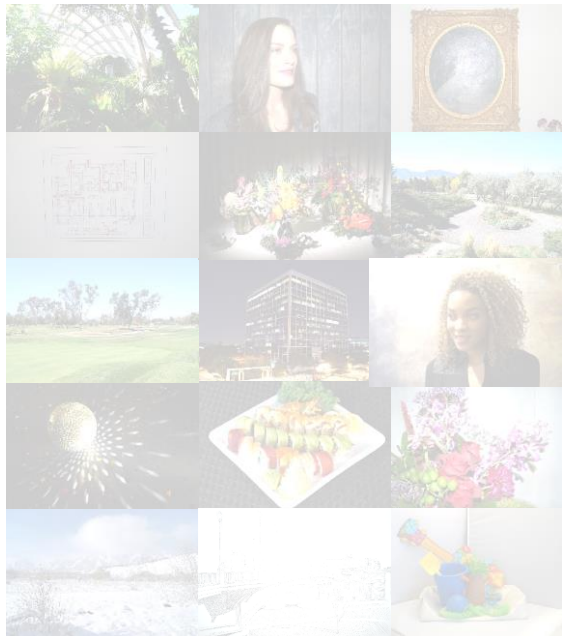

Assessing Video Quality



5 = Excellent
4 = Good
3 = Fair
2 = Poor
1 = Bad

Mean Opinion Score (MOS)

www.its.bldrdoc.gov



5 = Excellent
4 = Good
3 = Fair
2 = Poor
1 = Bad

Mean Opinion Score (MOS)

Metric Alternatives

Solutions

network performance
before / after compression
(only for pristine footage)

Technology Gap

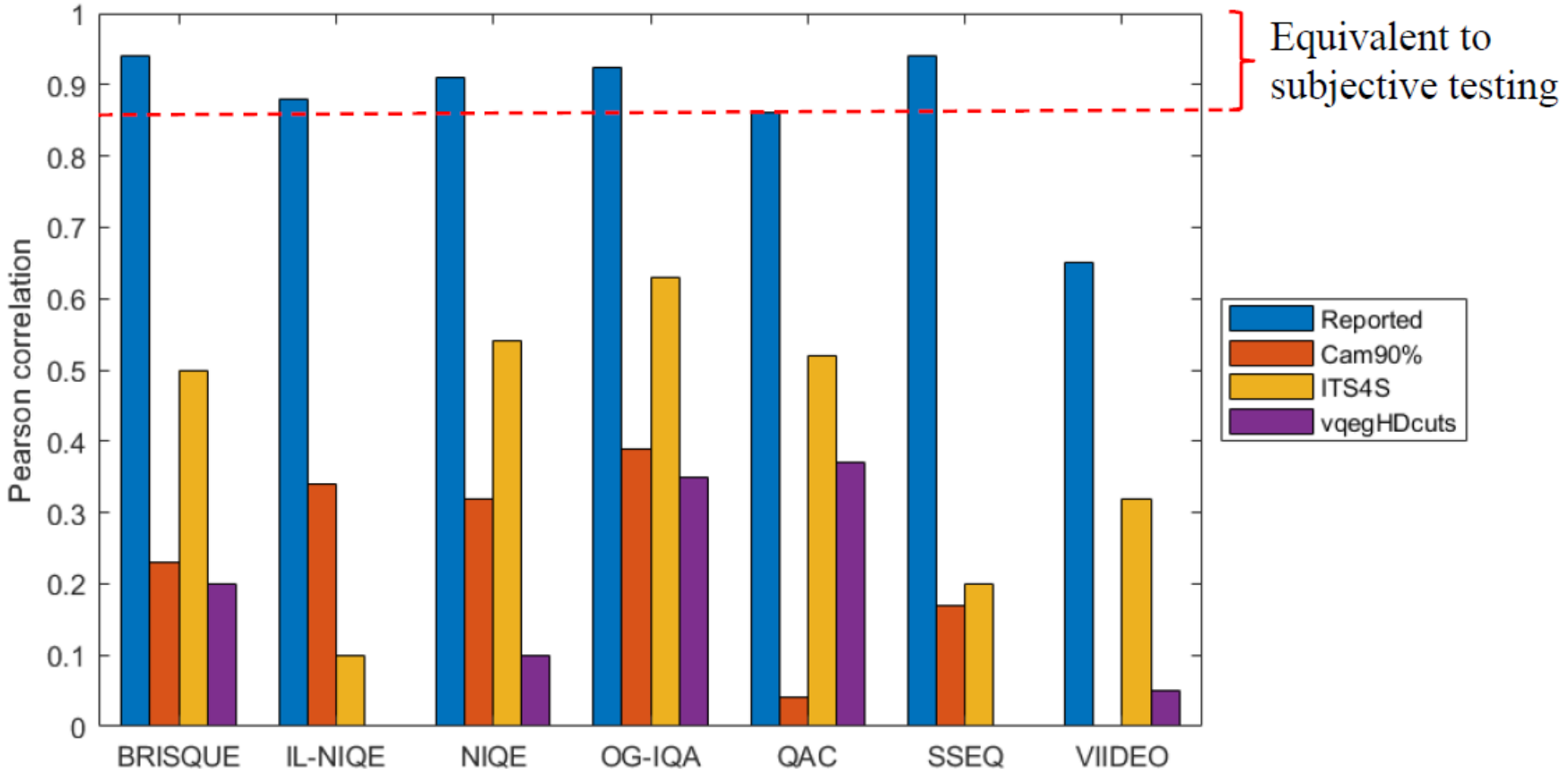
camera capture
user generated content

www.its.bldrdoc.gov

No Reference (NR) Metrics

- Analyze pixels (as displayed)
- Advantage: assess camera capture impairments
- Disadvantage: extremely difficult

Prior Work—Limited Scope, No Camera Impairments



Margaret H. Pinson, "Analysis of No-Reference Metrics for Image and Video Quality of Consumer Applications," NTIA Technical Memo TM-20-547, Jan. 2020

Inappropriate Training Data

Proprietary Algorithms

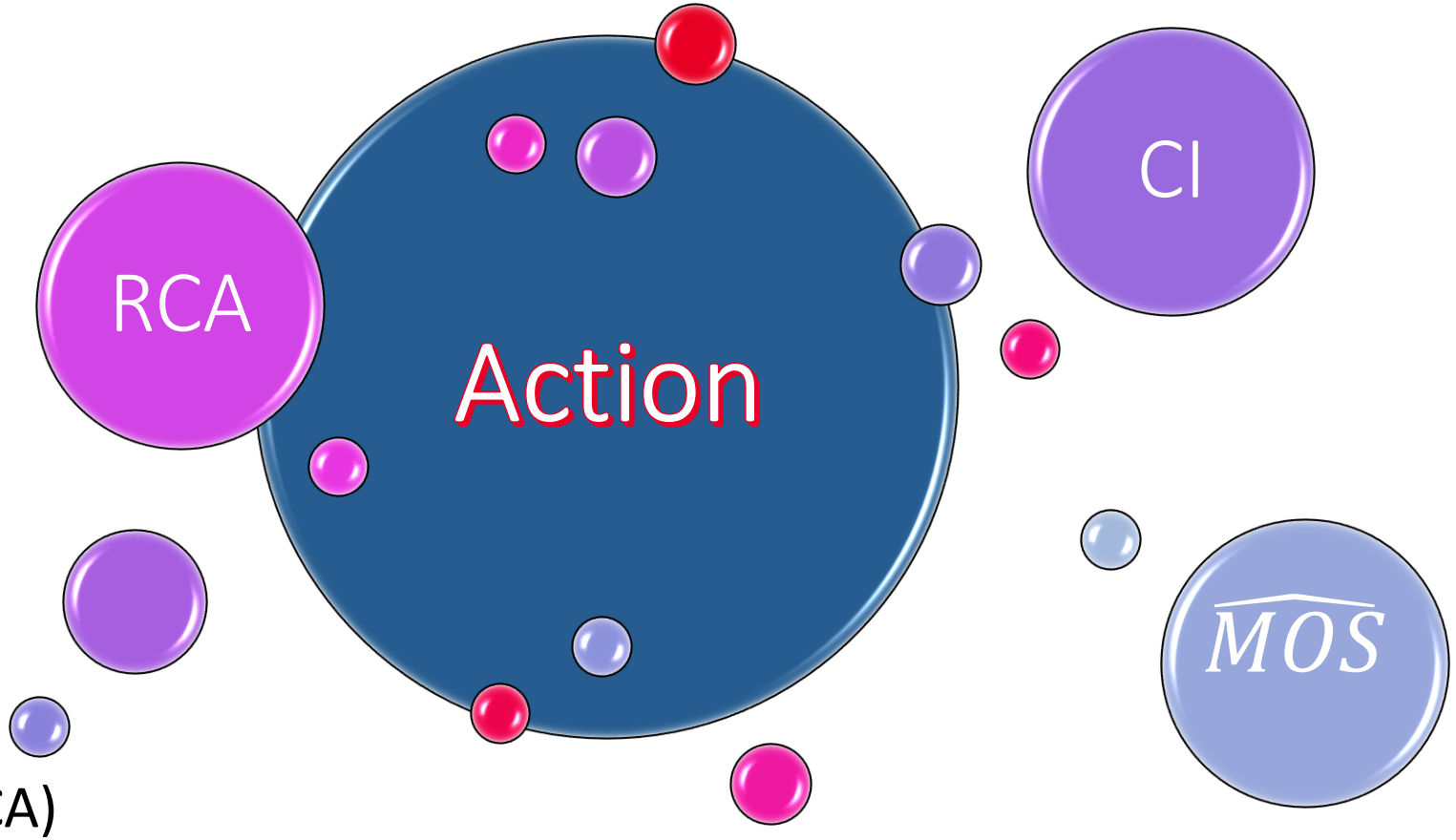
Wrong Question

Ephemeral Researchers

Research Tools

Research Paradigm

Metric Specifications



- Root cause analysis (RCA)
- Confidence interval (CI)
- Estimated mean opinion score ($\widehat{MOS}$)

Impairments

Spatial

Temporal

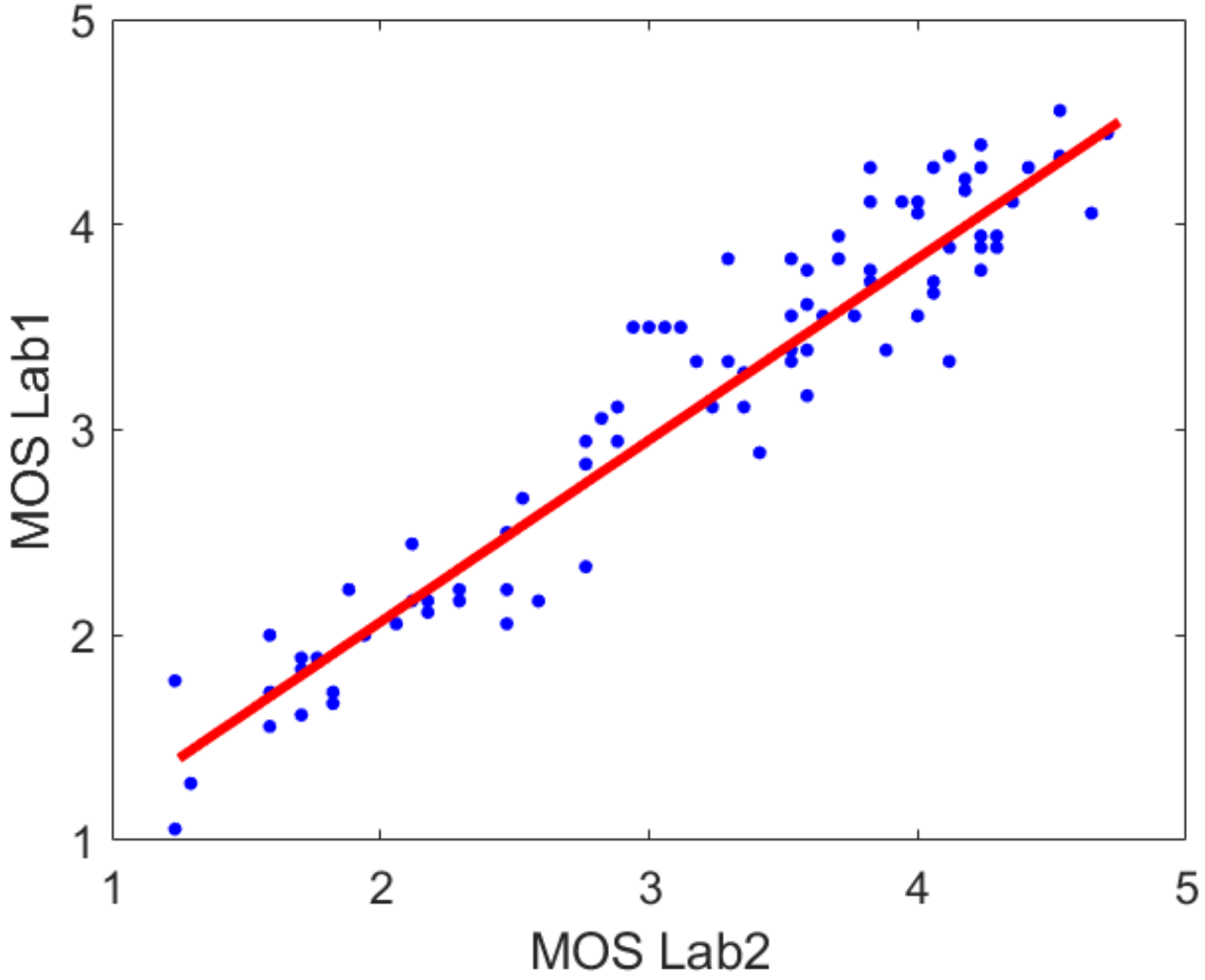
Research Paradigm — From RCA to $\widehat{MOS}$

- $\widehat{MOS} = 5 - \sum_{p=1}^{9}(w_p x_p)$
  - w_p weight for RCA parameter p
  - x_p value of RCA parameter p

- Modular programming
- Split the research effort by impairment
- Ignore impairment: w_p = 0

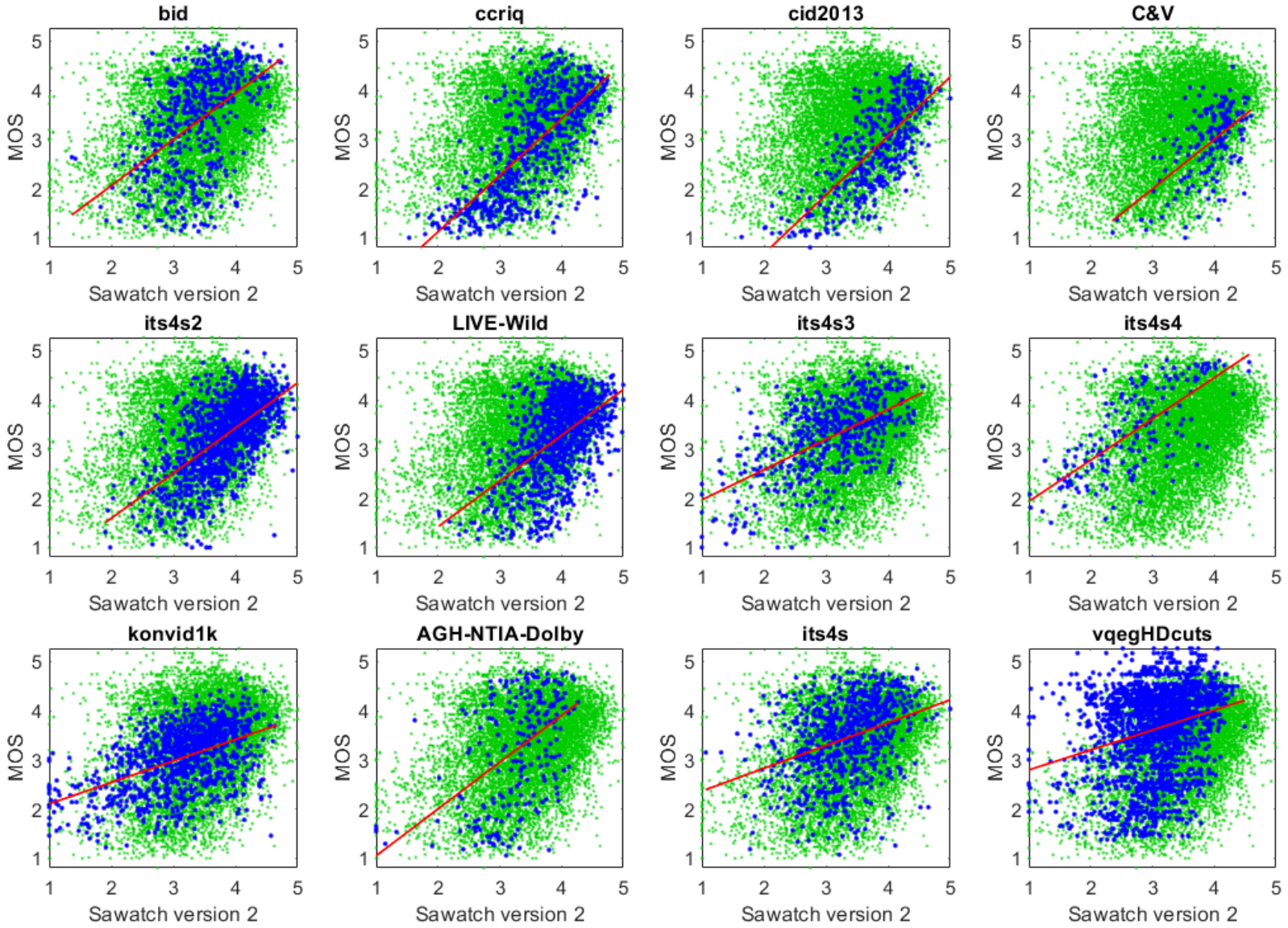
- Include
  - Image & video
  - Camera capture (1st)
  - Compression (2nd)
- Exclude
  - Temporal integration
  - Transmission errors

Scatter Plot for 2 Labs, Same Test



Sawatch Version 2

Image Dataset, Camera Capture	Pearson Correlation	False Ranking
BID	0.50	19%
CCRIQ	0.72	12%
CID2013	0.75	11%
CCRIQ2 & VIME1	0.58	15%
ITS4S2	0.63	12%
LIVE-Wild	0.53	—
Video Dataset, Camera Capture	**Pearson Correlation**	**False Ranking**
ITS4S3	0.58	15%
ITS4S4	0.72	11%
KonViD-1K	0.50	15%
Video Dataset, Compression	**Pearson Correlation**	**False Ranking**
AGH/NTIA/Dolby	0.50	21%
ITS4S	0.36	21%
vqegHDcuts	0.26	25%

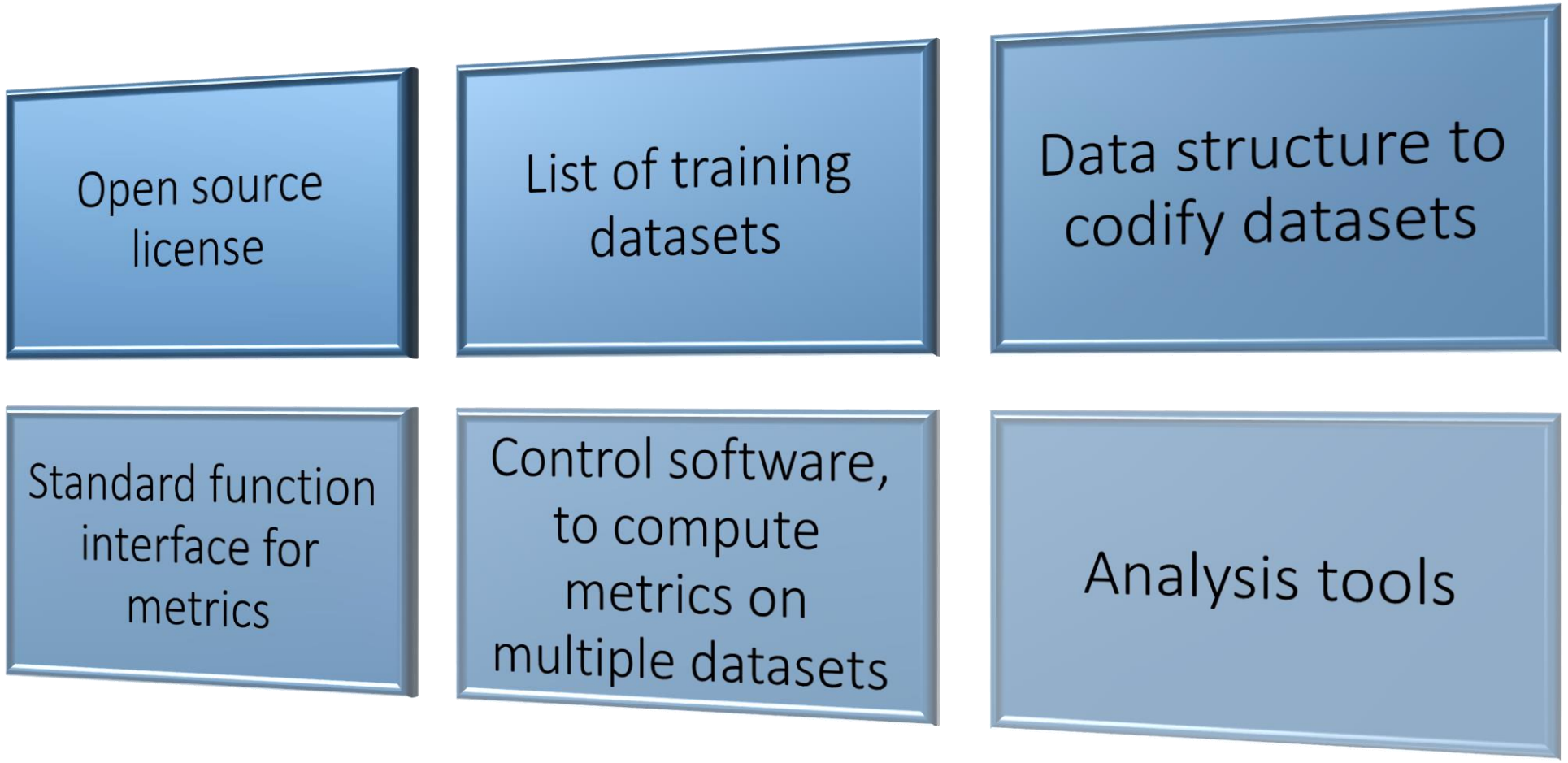


Sawatch Version 2 RCA

Goal	Metric Name	Rating
MOS	Sawatch version 2	☆ ☆ ☆
RCA	Blur	☆ ☆
RCA	Fine Detail	☆ ☆ ☆
RCA	White Level	☆ ☆ ☆
RCA	Black Level	☆ ☆
RCA	Pan Speed	☆ ☆
RCA	Color Noise	☆ ☆
RCA	Super Saturation	☆ ☆
RCA	Pallid	☆ ☆
RCA	Blockiness	☆ ☆

Rating Scale	Definition
☆	Very inaccurate
☆ ☆	Promising results
☆ ☆ ☆	Consistent performance
☆ ☆ ☆ ☆	One person ad-hoc test
☆ ☆ ☆ ☆ ☆	Six person pilot test

GitHub Repository – NRMetricFramework



Future Vision

MOS 3.5

RCA Interlacing artifacts

NR Metric Sawatch

- Download from GitHub
  - https://github.com/NTIA/NRMetricFramework/releases
  - Any purpose, commercial or non-commercial
  - Reports on other NR metrics
- Need
  - R&D videos with broadcast impairments
  - Can you Contribute?

- Margaret H Pinson, mpinson@ntia.gov