bit.ly/SMW_VQ

T101. HOW-TO: FINE-TUNING YOUR ADAPTIVE ENCODING GROUPS WITH OBJECTIVE QUALITY METRICS

Jan Ozer
www.streaminglearningcenter.com
jozer@mindspring.com/
276-235-8542
@janozer

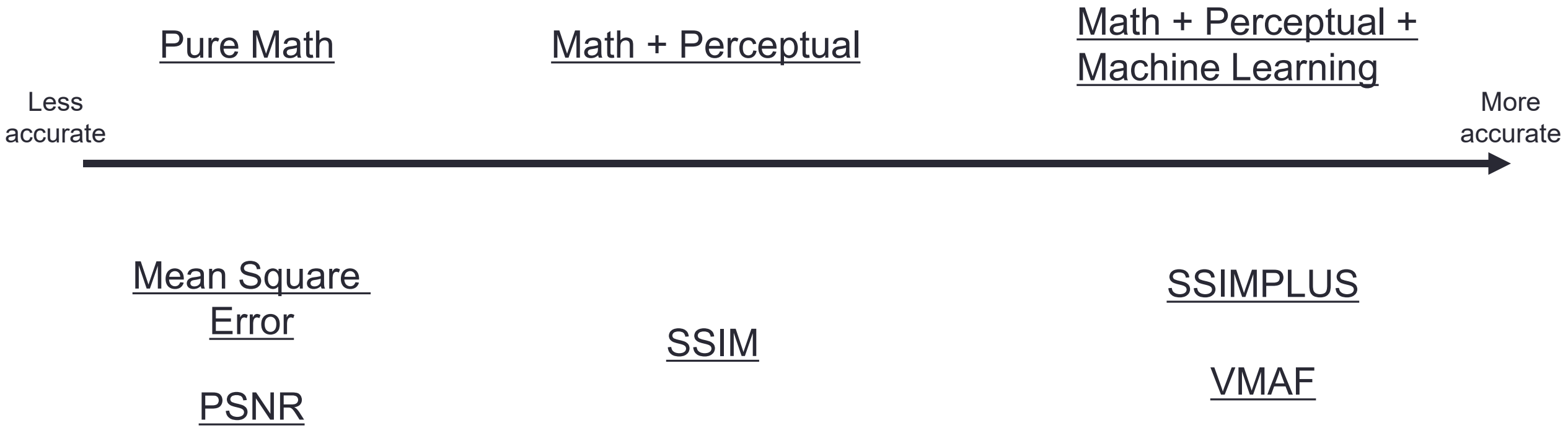
Agenda

- Overview of Objective Quality Metrics
- Tools/Metrics I use
- Building your encoding ladder
- Other configuration options

What Are Objective Quality Metrics

- Mathematical formulas that (attempt to) predict how human eyes would rate the videos
  - Faster and less expensive
  - Automatable
- Examples
  - MOS (Mean Opinion Score)
  - Peak Signal to Noise Ratio (PSNR)
  - Structural Similarity Index (SSIM)
  - SSIMPlus
  - VMAF (Video Multimethod Assessment Fusion)

Accuracy Continuum

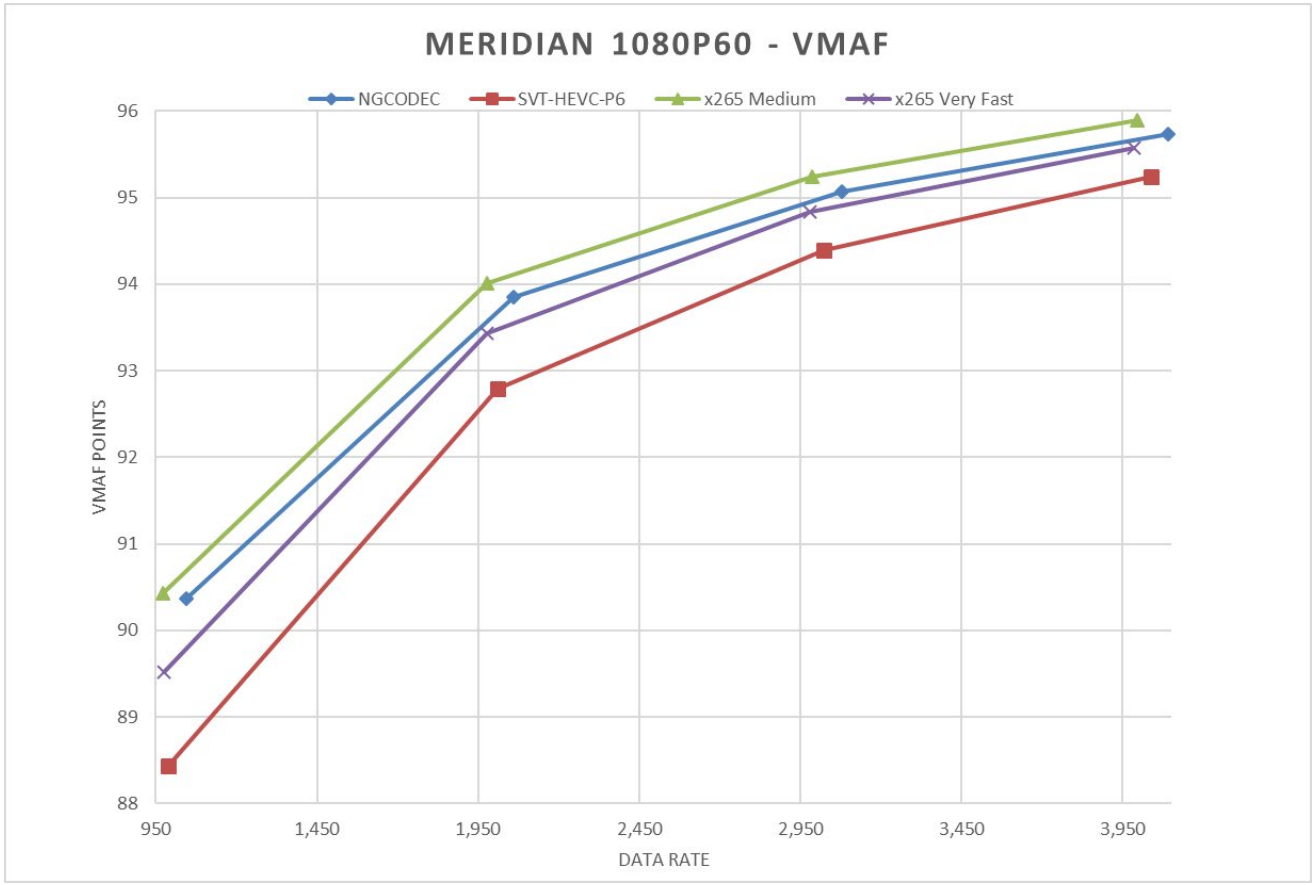


Tying Metrics to Predicted Subjective Ratings

	MOS	PSNR	SSIM	SSIMPLUS	VMAF
Scoring	1 – 5	0 – 100	0 – 1	0 - 100	0-100
No artifact threshold	NA	45 dB	.99	100	93
Artifacts present	NA	35 dB	.5	NA	NA
Interpreting scores					
Excellent	5	45+	.99 +	80 – 100	80 – 100
Good	4	38	.95 - .99	60 – 80	60 – 80
Fair	3	30	.88 - .98	40 – 60	40 – 60
Poor	2	24	.50-.88	20 – 40	20 – 40
Bad	1	< 15	< .5	< 20	< 20
Just noticeable difference	NA	NA	NA	NA	6
Device ratings	No	No	No	Multiple	Standard, Phone, 4K
Ownership	Open source	Open source	Open source	Proprietary	Open Source

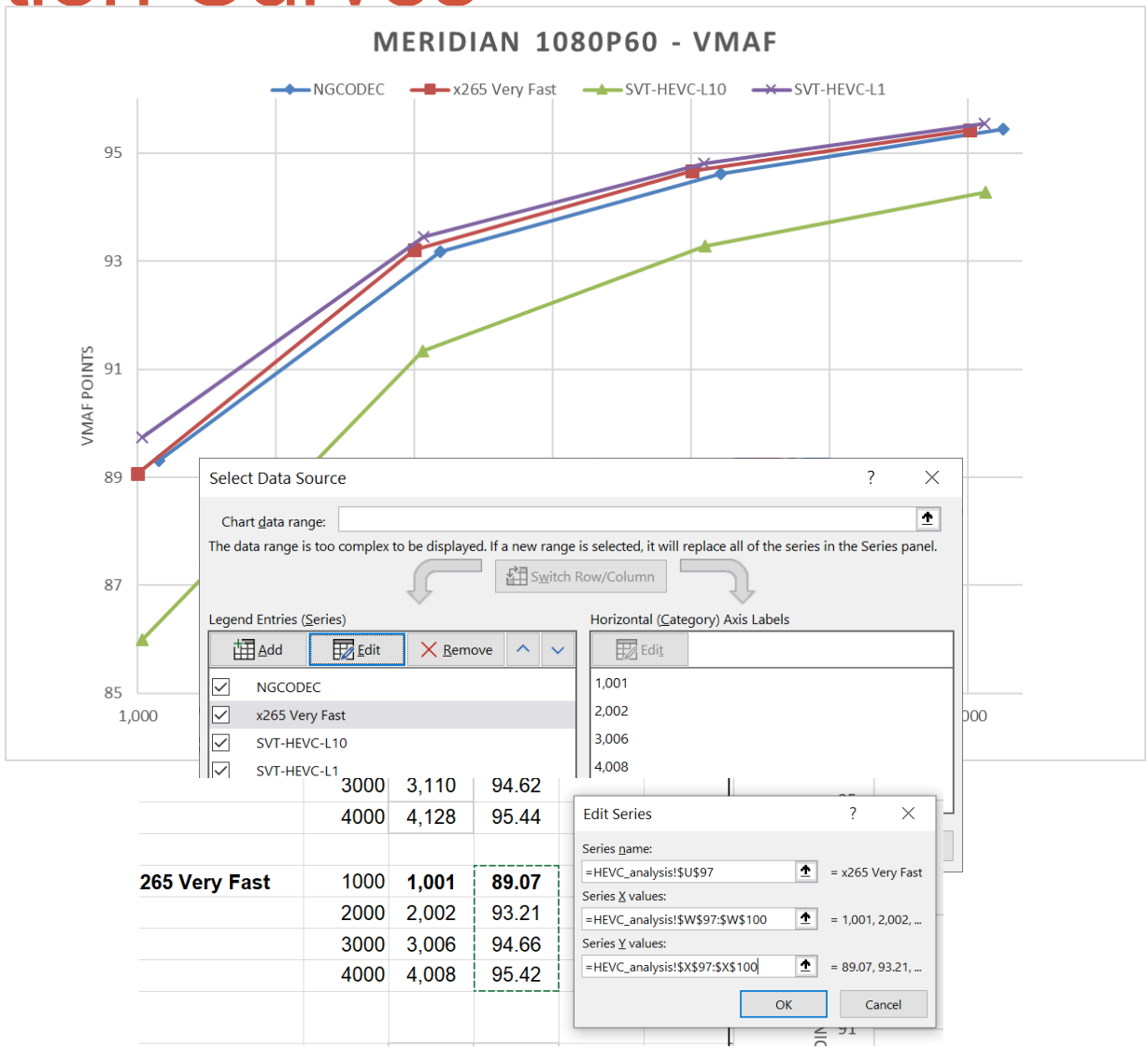
Key Concepts: Rate Distortion Curves and BD-Rate Functions

- Formal numbers-only analysis, typically deployed for codec/encoder comparisons
- Rate-distortion curve
  - Four encodes with different technologies (VMAF)
    - On right – HEVC transcoders for live broadcasts
  - Rate-distortion curve – how each technology “distorts” at the various data rates



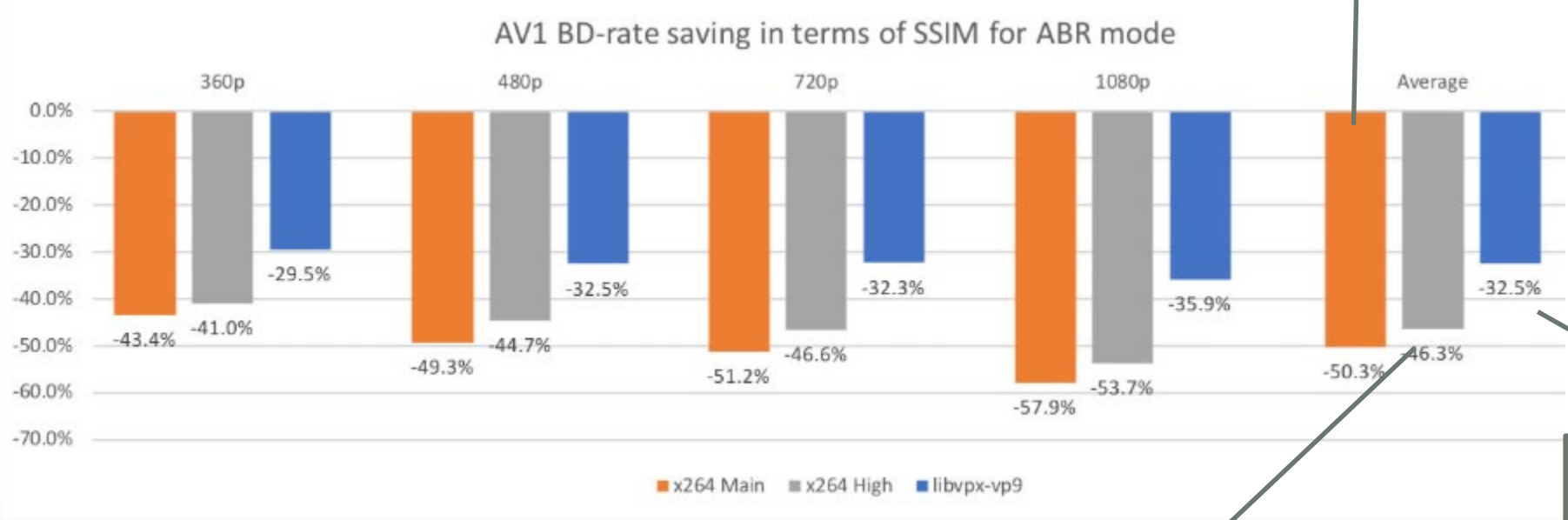
Visualization – Rate Distortion Curves

- Overview
  - Can do in Sheets but Excel clearer and simpler
- Format data
- Create chart
  - Must be Scatter with straight lines and markers
- Insert data
- Customize graph area
- Rinse and repeat



Facebook Analyzes AV1

AV1 produced same quality as x264 main profile at 50.3% lower data rate

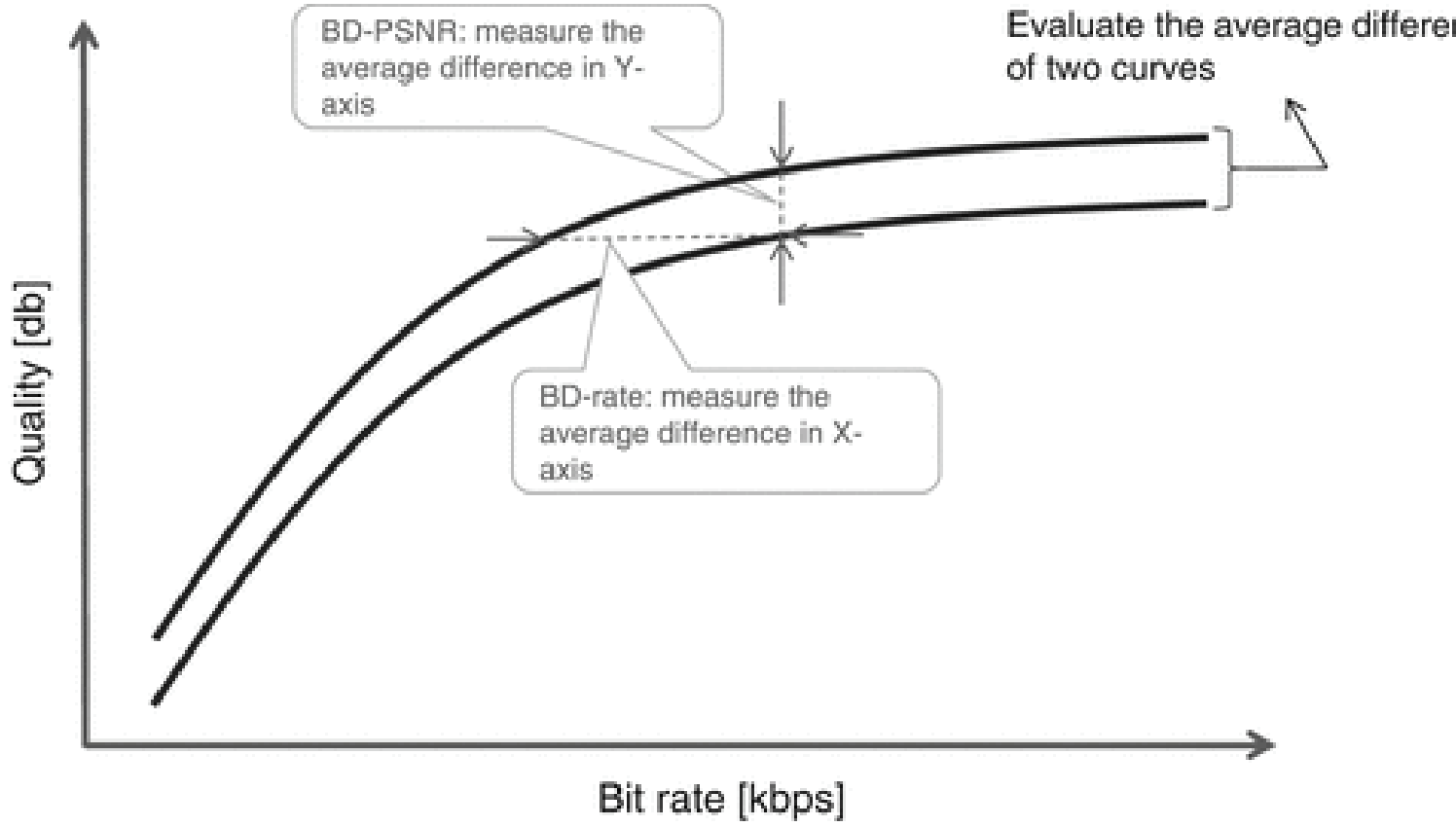


AV1 produced same quality as x264 high profile at 46.3 % lower data rate

AV1 produced same quality as VP9 at 32.5% lower data rate

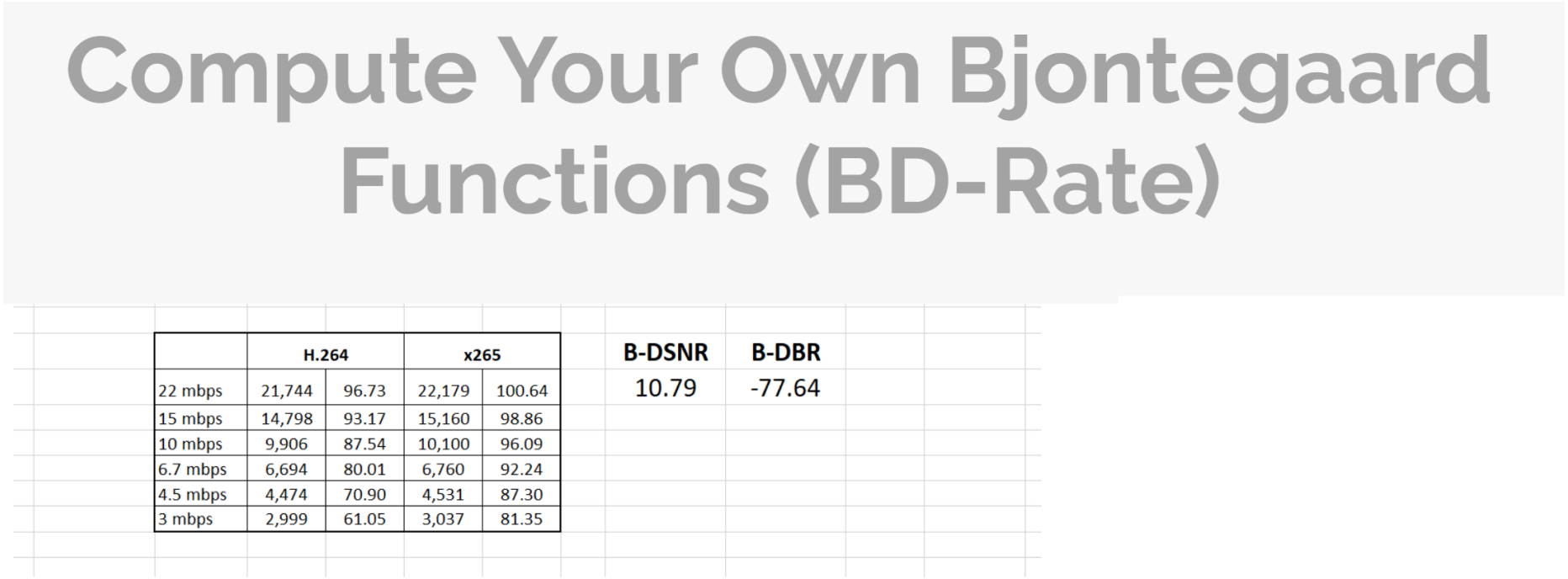
Bjontegaard Functions

- Quantifies differences between two curves
  - BD-Rate – data rate saving for the **same quality**
  - BD-PSRN – quality disparity for same **bitrate**
    - Can use with any metric



http://bit.ly/BDRPSNR

Download Excel Spreadsheet with Macro



	H.264		x265	
22 mbps	21,744	96.73	22,179	100.64
15 mbps	14,798	93.17	15,160	98.86
10 mbps	9,906	87.54	10,100	96.09
6.7 mbps	6,694	80.01	6,760	92.24
4.5 mbps	4,474	70.90	4,531	87.30
3 mbps	2,999	61.05	3,037	81.35

B-DSNR	B-DBR
10.79	-77.64

- http://bit.ly/BD_functions

Macro – BD-RATE

- Always referential and have to pick the reference
  - Here, SVT is reference
  - Result – On average, NGCodec can produce same quality as SVT at data rate reduction of 4.21%
- BD-BR macro
  - Blue – bitrate of reference file (SVT)
  - Red – metric score of reference file (SVT)
  - Purple – bitrate of target file (NGCodec)
  - Green – metric score of target file (NGCodec)

This value

NGCODEC				SVT-HEVC-P6		
Dinnerscene - 1080p	Bitrate	VMAF	BDRate	Dinnerscene -	Bitrate	VMAF
1000	1,029	80.20	-4.21	1000	970.1	79.22
2000	2,029	87.68		2000	1971	87.00
3000	3,029	90.75	BD Quality	3000	2970	90.17
4000	4,029	92.48	0.39	4000	3970	91.94

NGCODEC				SVT-HEVC-P6		
Dinnerscene - 1080p	Bitrate	VMAF	BDRate	Dinnerscene -	Bitrate	VMAF
1000	=BDBR(AF3:AF6,AG3:AG6,$AB3:$AB6,$AC3:$AC6)					79.22
2000	2,029	87.68		2000	1971	87.00
3000	3,029	90.75	BD Quality	3000	2970	90.17
4000	4,029	92.48	0.39	4000	3970	91.94

Computed with similar macro

Computed with this macro

The Tools I Use or Recommend

- Moscow State University Visual Quality Comparison Tool (VQMT)
- SSIMPLUS VOD Monitor
- Hybrik (Dolby) Media Analyzer

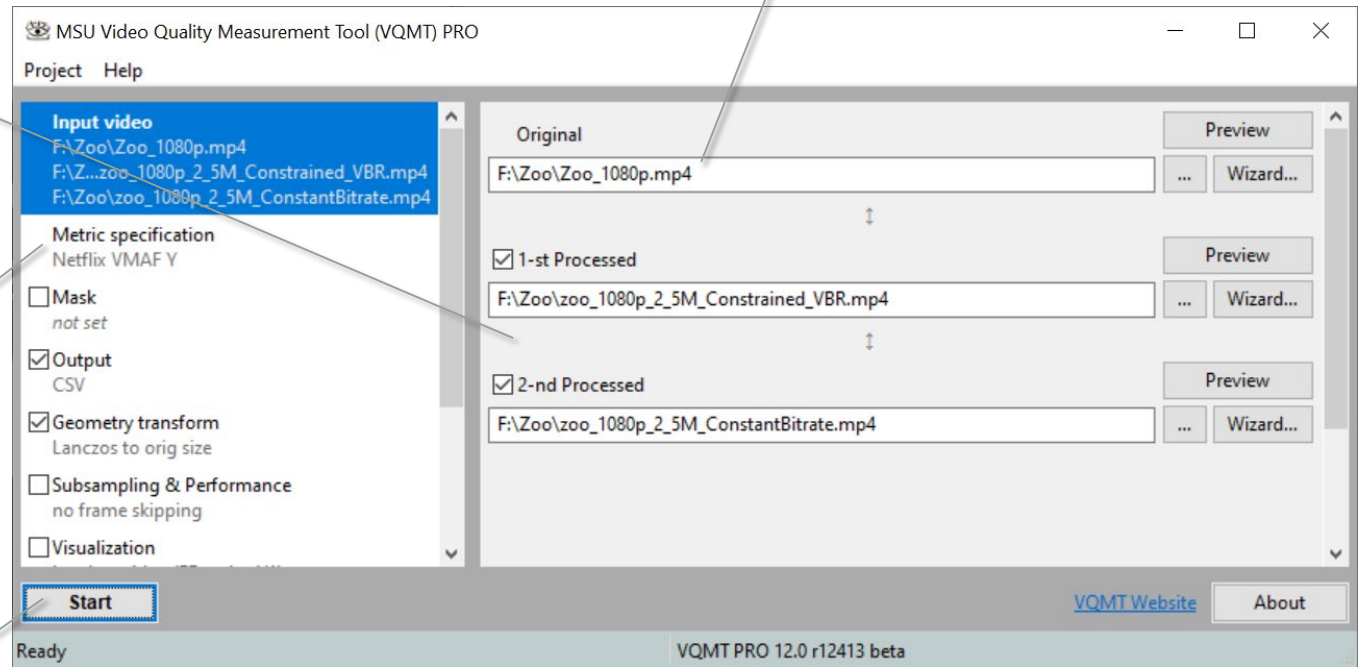
VQMT Workflow

Load Source File

Load one or two encoded files

Choose Metric

Press Start



Results Visualization

Score entire comparison

Orange– first file
Green– second

Zoom in of selection

Slide through frames

Click to Show Actual Frames

See Frames

VBR frame = 69.24
CBR frame =47.73

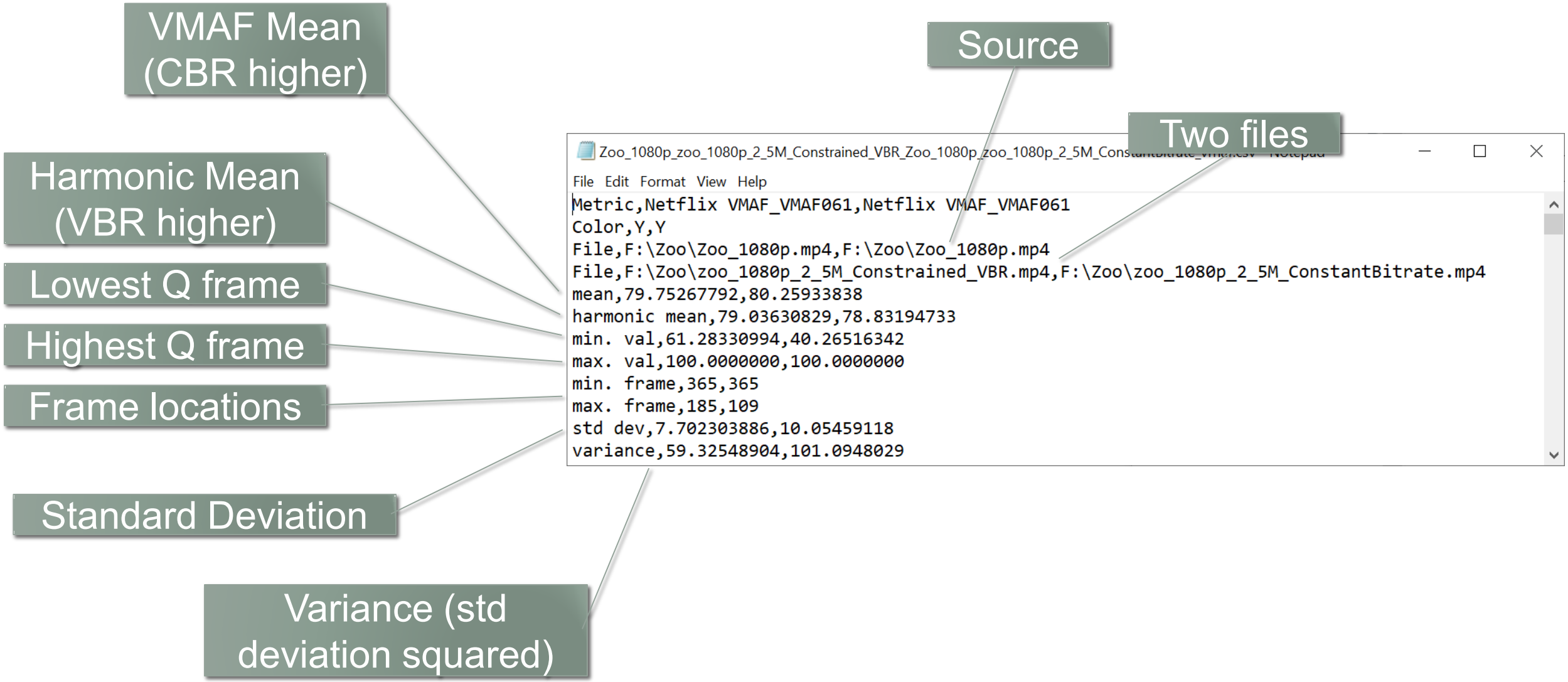
VBR-encoded

- Slide through video file
- Compare

CBR-encoded

- Compare side-by-side or hot key between original and two encoded files

Numerical Results from CSV



MSU VQMT

Pros

- Affordable (~$995)
- GUI and command line
- Very visual – easy to see test results in actual frames
- Multiple algorithms – VMAF, PSNR, SSIM, MS SSIM
- My review of VQMT
  - bit.ly/VQMT_review

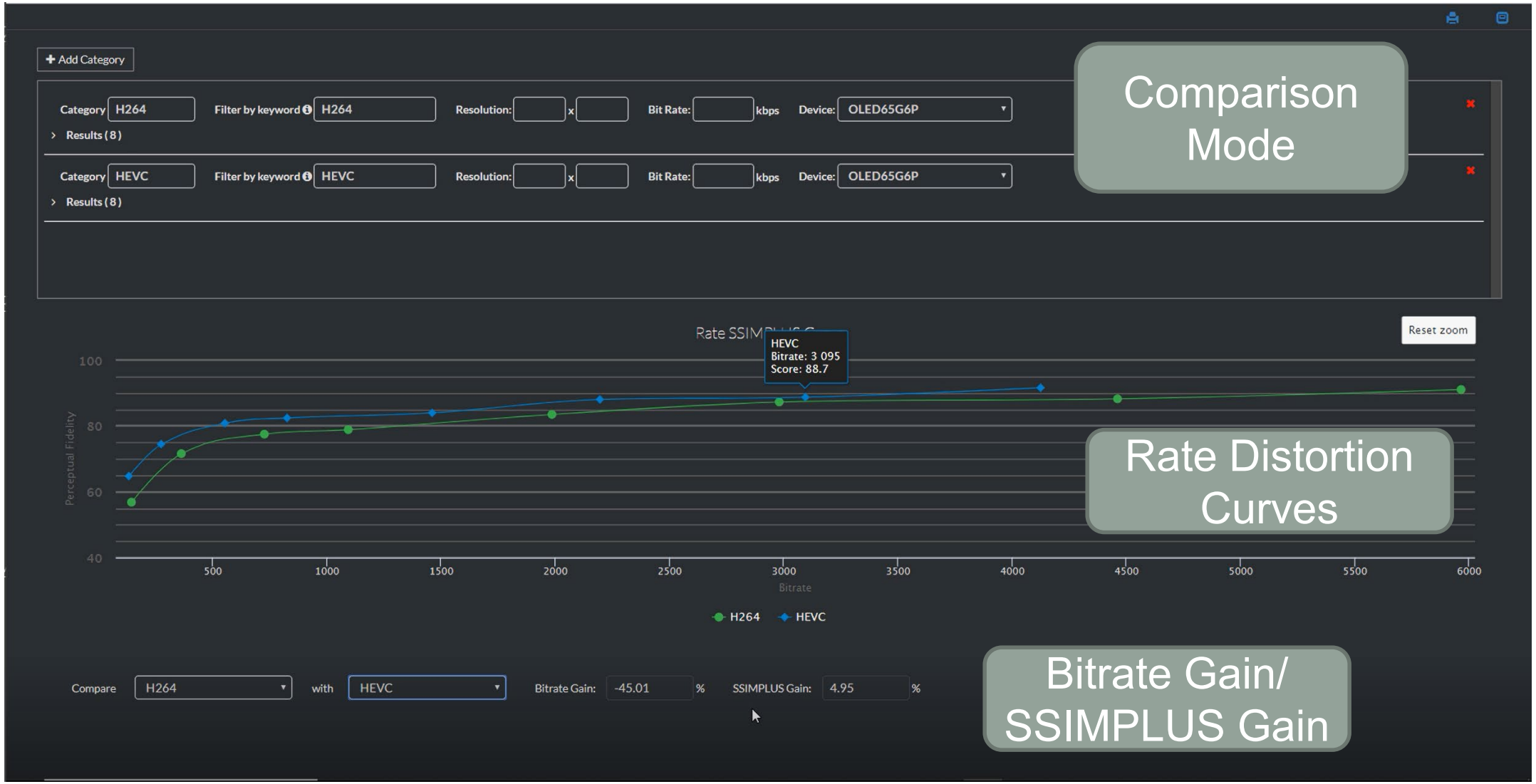
Cons

- Can only compare files of:
  - Like frame rate
- Time consuming data entry
  - Can output JSON for automated input

SSIMWAVE VOD Monitor

- Based on SSIMPLUS Algorithm
- Rates videos on scale that corresponds with human perception
  - 80 – 100 – Excellent
  - 60 – 80 – Good
  - 40 – 60 – Fair
  - 20 – 40 – Poor
  - < 20 – Bad
- Predicts ratings on multiple devices
  - Phones, TVs, monitors, etc.
- Can compare different resolutions (without conversion)
- Can compare different frame rates (without conversion)
- Here at Streaming Media West

SSIMPLUS VOD Monitor



Other Tools

- Hybrik – part of Dolby family
  - Cloud system – I used for large comparisons – process hundreds of files overnight
  - Not available as analysis-only
  - Cost structure is minimum $1,000/month
- FFmpeg/VMAF Master
  - Cheap and accessible
  - Number only; no visualization

- Installing and Using Netflix VMAF-Master, Streaming Learning Center, March 2019 (http://bit.ly/VMaster)
- Compute VMAF Using FFmpeg on Windows, Streaming Learning Center, November 2019 (bit.ly/ff_vmaf_win)

Applying Metrics in Technology Decision Making

- Start with a number
- Check results plot for transient issues or low frame values
- Check the actual frames
- If visible, check video

Start with the Number

- Checking the difference between CBR and VBR (both 1080p@2500)
  - 200% constrained VBR - 80.47
  - 1-pass CBR - 79.97
- Both very good, 1-pass CBR cuts encoding time in half, let's use that!

But Wait – Let's Look at Results Plot

But Wait – Let's Look at Results Plot

- Video removed to reduce file size

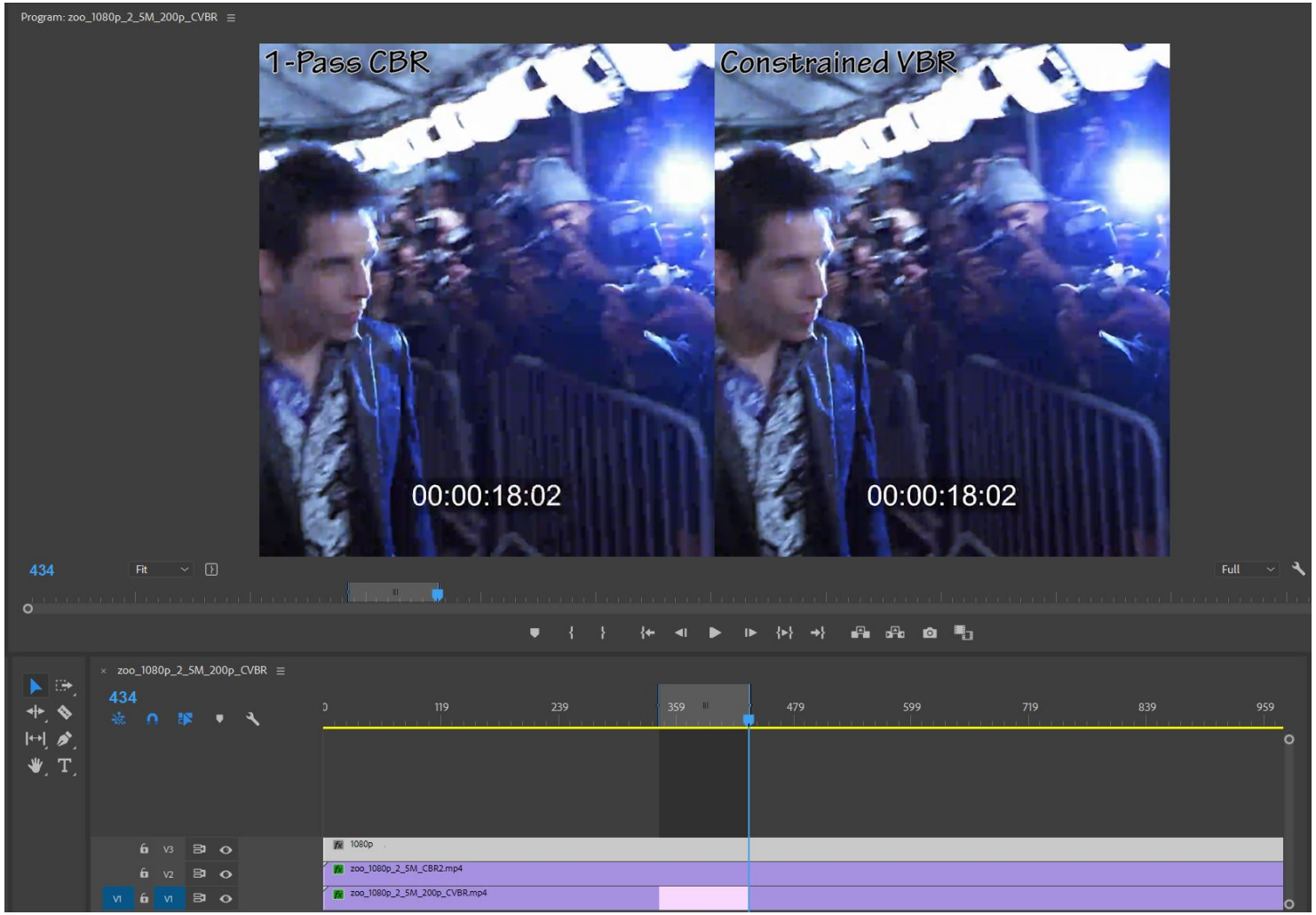
Let’s Look at Frames - Original

Let’s Look at Frames – Constrained VBR

Let’s Look at Frames - CBR

But Can You See the Difference In Real Time?

Load Files into Video Editor



Load Files into Video Editor

- Video removed to reduce file size

SSIMPLUS VOD Monitor

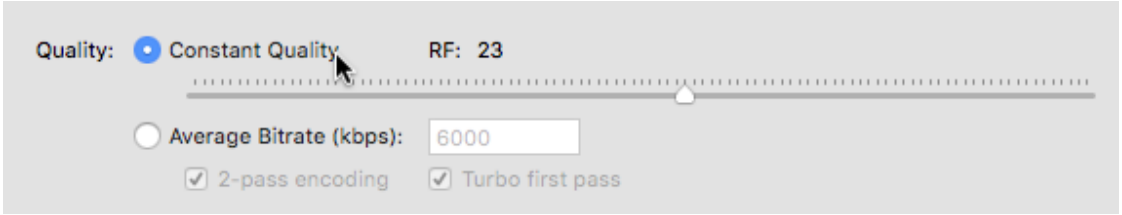
- Video removed to reduce file size

Using Quality Metrics – Finding the Top Rung

- About CRF
- Configuring ladder
  - Floor and ceiling

About Constant Rate Factor Encoding

- Encoding mode available in x264, x265, VP8/9
- Encodes to a specific *quality* level, not a data rate
- Two uses
  - As gauge of encoding complexity
  - With caps, a per-title encoding technique



- Range is 1-51
  - Lower number means higher quality
  - For 2D video, CRF 23 roughly delivers VMAF 93

Finding the Optimal Data Rate (Per-title)

- Compute data rate with CRF 23
  - Values varied from 1,001 to 6,111 (over 600%)
- Measure VMAF rating
  - Values ranged from 92.74 to 96.88
  - Standard deviation was 1.39 (pretty small)
- Analysis
  - At 2.7 Mbps, a talking head video offers same quality as movie at 6.1 Mbps (even more for synthetic videos)
  - Validating the benefits of per-title encoding

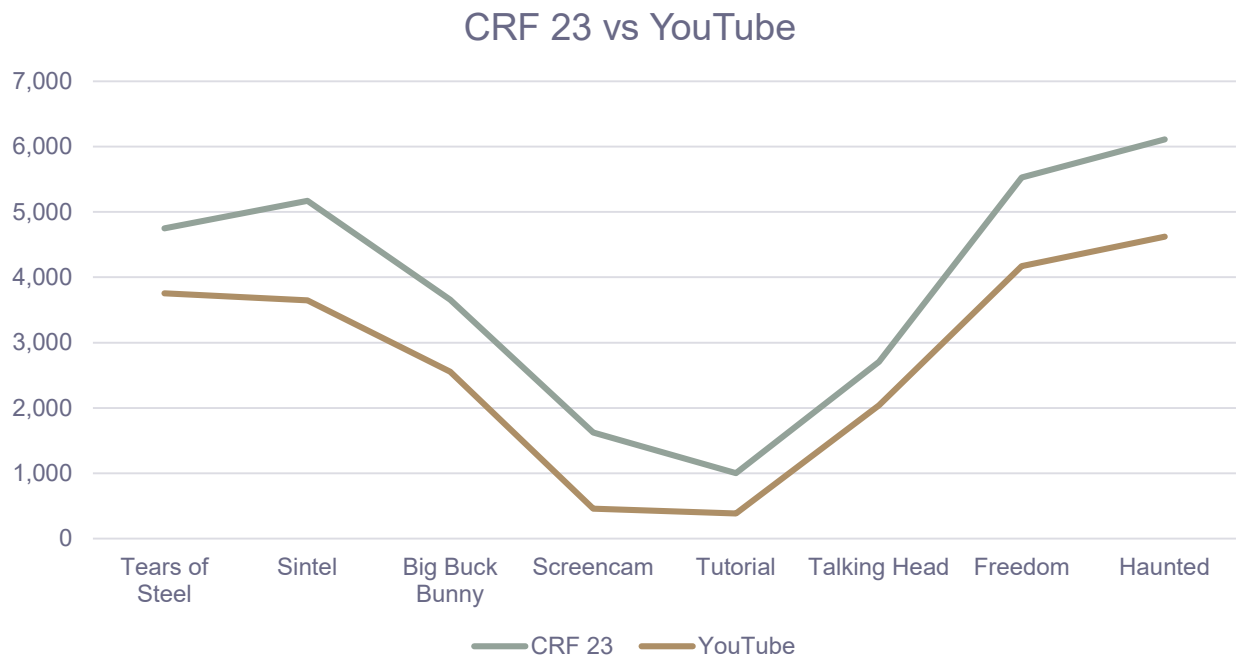
CRF23 - 1080p	FPS	Description	Data Rate	VMAF
Tears of Steel	24	Real world/CG movie	4,747	96.45
Sintel	24	Complex animation	5,168	96.96
Big Buck Bunny	30	Simple animation	3,657	96.88
Screencam	30	Camtasia-based video	1,625	96.59
Tutorial	30	PowerPoint and talking head	1,001	96.68
Talking Head	30	Simple talking head	2,706	95.47
Freedom	30	Concert footage	5,527	95.90
Haunted	30	DSLR movie-like production	6,111	92.74
Average			**3,818**	**95.96**
Standard deviation				**1.39**

- Conclusion:
  - CRF maps accurately to VMAF values

VMAF Verification – 93 is the Number

- Real Networks White Paper - VMAF Reproducibility: Validating a Perceptual Practical Video Quality Metric
  - 4K 2D videos
- ***VMAF score of about 93 = video that is either indistinguishable from original or with noticeable but not annoying distortion.***
  - http://bit.ly/vrqm_5

CRF 23 Reality Check: YouTube Comparison



- Upload files to YouTube; measure data rate
- YouTube uses AI-based per-title optimization
- Pattern very similar
- YouTube averages 1 Mbps lower
- 3 VMAF points lower (1/2 JND)

So

- 1080p videos, CRF 23 = ~93 VMAF = shippable quality
- Significant data point
  - As we'll see – encoding ladder starts at the top

Once You Have Highest it Becomes Math Exercise

- Step 1: Choose highest
- Step 2: Choose lowest
- Step 4: fill in the blanks (between 150/200% apart)

200 kbps

500 kbps

1000 kbps

1600 kbps

2100 kbps

3100 kbps

4600 kbps

Then Question is:

- Netflix approach
  - Compute VMAF scores at multiple resolutions at each data rate
  - Choose best quality at each resolution
  - VMAF proven for 2D by Netflix, what about 3D?

Zap1 - VMAF	4K	2K	1080p	720p	480p	360p	240p
5000	90.19	89.70	84.82				
4500	89.58	88.23	84.38				
4000	88.43	87.50	83.84				
3800	87.88	87.14	83.58				
3600	87.27	86.71	83.25				
3400	86.60	85.72	82.87				
3200	85.80	85.40	82.45				
3000	85.03	85.09	82.01				
2800	83.97	84.34	81.43				
2600	82.86	83.50	80.85				
2400	81.45	82.51	80.09	71.92			
2200	79.79	81.24	79.20	71.35			
2000	77.94	79.82	78.04	70.66			
1800		78.11	76.73	69.70	53.28		
1600		75.91	74.93	68.41	52.82		
1400		73.26	72.64	66.89	52.13	32.07	
1200		69.83	69.69	64.68	51.05	31.75	
1000		65.15	65.75	61.64	49.36	31.17	
900		62.26	63.25	59.64	48.11	30.76	
800		58.69	60.27	57.20	46.54		
700		54.29	56.62	54.06	44.63	29.18	
600		48.79	52.32	50.65	42.02	27.84	
500			46.65	45.96	38.74	25.92	
400			39.06	40.23	34.21	23.11	
300			28.52	32.68	x		
200			13.88	21.73	x		

How Ladders Change for Advanced Codecs

- Need completely different ladder for HEVC/VP9/AV1

Proof – Tears of Steel

H.264

H.264	1080p	720p	540p	432p	360p	270p	234p
5000	96.22						
4800	96.01						
4600	95.80	95.27					
4400	95.55	95.10					
4200	95.30	94.96					
4000	94.96	94.73					
3800	94.60	94.53					
3600	94.14	94.30					
3400	93.70	93.99					
3200	93.11	93.64					
3000	92.48	93.24					
2800	91.70	92.78					
2600	90.75	92.25					
2400	89.70	91.59	90.39				
2200	88.37	90.80	89.76				
2000	86.72	89.85	88.95	86.93			
1800	84.68	88.66	88.00	86.10			
1600	82.13	87.13	86.77	85.02	81.58		
1400	78.65	85.19	85.16	83.67	80.28		
1200	73.91	82.56	83.01	81.84	78.57		
1000	67.39	78.86	80.02	79.24	76.19		
900	63.18	76.39	77.98	77.47	74.60	66.66	60.58
800	57.93	73.25	75.51	75.34	72.68	65.11	59.23
700	51.47	69.42	72.34	72.59	70.23	63.14	57.49
600	43.12	64.52	68.37	69.11	67.12	60.70	55.33
500	33.31	58.05	63.13	64.66	63.04	57.52	52.46
400	20.82	49.48	56.00	58.46	57.48	53.13	48.59
300	9.74	37.56	45.95	49.62	49.60	46.80	42.96
200	3.73	20.40	30.87	36.12	37.48	36.88	34.03
100		2.75	8.08	14.45	17.50	19.85	18.66

HEVC

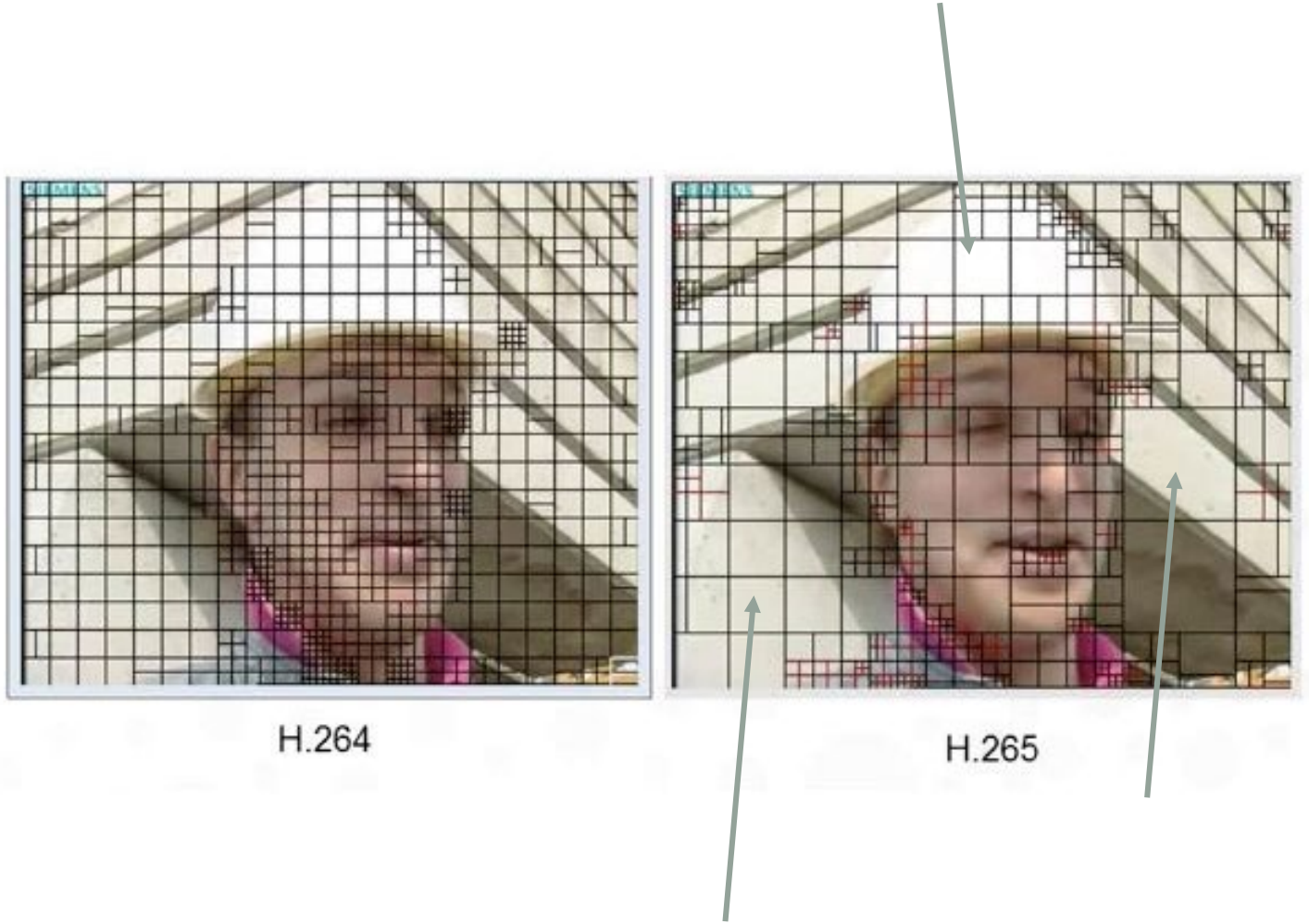
HEVC	1080p	720p	540p	432p	360p	270p	234p
5000	97.67						
4800	97.55						
4600	97.44						
4400	97.31						
4200	97.17						
4000	97.01						
3800	96.84						
3600	96.63						
3400	96.41						
3200	96.15	95.41					
3000	95.86	95.16					
2800	95.52	94.87					
2600	95.09	94.52					
2400	94.58	94.12	92.09				
2200	93.97	93.63	91.62				
2000	93.16	92	91.05	88.30			
1800	92.18	92.25	90.34	87.63			
1600	90.94	91.27	89.44	86.78	83.18		
1400	89.36	89.97	88.27	85.69	82.12		
1200	87.30	88.26	86.68	84.22	80.73		
1000	84.42	85.84	84.46	82.20	78.79		
900	82.39	84.21	83.02	80.86	77.51	68.45	62.18
800	80.03	82.20	81.23	79.19	75.91	67.09	60.92
700	77.04	79.67	78.90	77.07	73.91	65.38	59.35
600	73.10	76.34	75.88	74.29	71.36	63.21	57.34
500	68.11	71.98	71.82	70.61	67.89	60.30	54.69
400	61.01	65.92	66.31	65.54	63.19	56.29	51
300	50.13	57.34	58.21	58.06	56.18	50	45.69
200	25.00	44.30	45.88	46.47	45	40.96	37.13
100	4.14	13.75	24.62	26.16	25.85	23.86	21.53

1080p best quality at far lower data rates than H.264

Lower resolutions don't provide the best quality

Why is HEVC More Efficient?

- Simply a better codec
- One prominent advantage – larger block sizes
  - H.264 – 16x16
  - HEVC – 64x64
- Can encode large images more efficiently



What About Different Types of Content?

- In general:
  - Synthetic videos encode at higher quality at lower bitrates (not shown here)
  - Look better at higher resolutions
    - Push 1080p lower down in the encoding ladder
    - Push 720p further down the ladder
- Not huge difference here, but much more profound for screencams and similar videos

Tears of Steel (real world/CG)

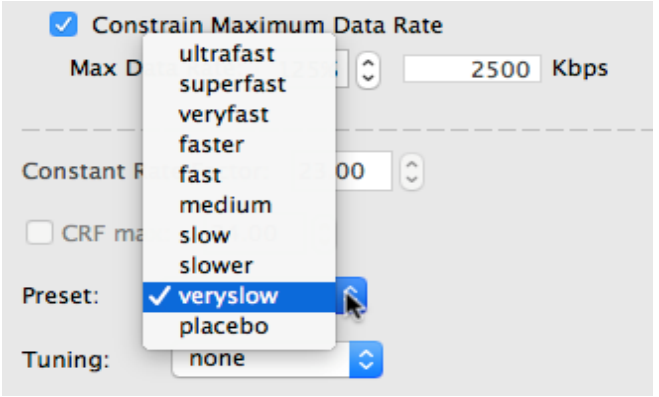
HEVC	1080p	720p	540p	432p	360p	270p	234p
5000	97.67						
4800	97.55						
4600	97.44						
4400	97.31						
4200	97.17						
4000	97.01						
3800	96.84						
3600	96.63						
3400	96.41						
3200	96.15	95.41					
3000	95.86	95.16					
2800	95.52	94.87					
2600	95.09	94.52					
2400	94.58	94.12	92.09				
2200	93.97	93.63	91.62				
2000	93.16	93.02	91.05	88.30			
1800	92.18	92.25	90.34	87.63			
1600	90.94	91.27	89.44	86.78	83.18		
1400	89.36	89.97	88.27	85.69	82.12		
1200	87.30	88.26	86.68	84.22	80.73		
1000	84.42	85.84	84.46	82.20	78.79		
900	82.39	84.21	83.02	80.86	77.51	68.45	62.18
800	80.03	82.20	81.23	79.19	75.91	67.09	60.92
700	77.04	79.67	78.90	77.07	73.91	65.38	59.35
600	73.10	76.34	75.88	74.29	71.36	63.21	57.34
500	68.11	71.98	71.02	70.61	67.89	60.30	54.69
400	61.01	65.92	66.31	65.54	63.19	56.29	51.05
300	50.13	57.34	58.21	58.06	56.18	50.40	45.89
200	25.00	44.30	45.88	46.47	45.24	40.96	37.13
100	4.14	13.75	24.62	26.16	25.85	23.86	21.53

Sintel (animation)

HEVC	1080p	720p	540p	432p	360p	270p	234p
5000	97.83						
4800	97.74						
4600	97.63						
4400	97.50						
4200	97.36						
4000	97.19						
3800	97.01						
3600	96.78						
3400	96.52						
3200	96.22	94.39					
3000	95.86	94.11					
2800	95.45	93.78					
2600	94.94	93.40					
2400	94.32	92.93	89.84				
2200	93.62	92.37	89.34				
2000	92.72	91.69	88.71	85.40			
1800	91.63	90.84	87.94	84.72			
1600	90.21	89.76	87.00	83.84	79.64		
1400	88.44	88.36	86.74	82.74	78.62		
1200	86.02	86.39	84.07	81.24	77.24		
1000	82.81	83.73	81.70	79.13	75.35		
900	80.79	82.02	80.16	77.76	74.10	64.67	58.74
800	78.22	79.83	78.25	76.06	72.55	63.43	57.63
700	75.22	77.22	75.91	73.94	70.64	61.88	56.22
600	71.44	73.84	72.94	71.27	68.17	59.87	54.42
500	66.61	69.68	69.13	67.71	64.90	57.24	52.02
400	60.19	63.98	63.94	62.97	60.47	53.61	48.73
300	48.81	56.19	56.62	56.22	54.16	48.26	43.81
200	26.36	44.11	45.66	53.05	44.22	39.79	36.06
100	5.17	15.45	23.86	26.96	26.53	24.50	21.89

Choosing the Optimal Encoding Time/Quality Tradeoff

- All encoders/codecs have configuration option that trades off time vs. quality
  - This technique lets you choose the best option
- Here – looking at x264 presets. What are presets?
  - Simple way to adjust multiple parameters to trade off encoding speed vs. quality
  - Used by virtually all x264 encoders
  - Medium is generally the default preset



When to Use This Technique

- When evaluating new encoders
- When choosing/evaluating encoding settings
- When comparing codecs

Test Procedure

- Choose test files
  - 1 movie (Tears of Steel)
  - 2 animations (Sintel, BBB)
  - Two general purpose (concert, advertisement)
  - One talking head
  - Screencam
  - Tutorial (PPT/Video)
- 2. Encode to all presets targeting around 96 VMAF max
  - All files encoded to different bitrates
- 3. Measure encoding time
- 4. Measure Average VMAF
- 5. Measure Low-Frame VMAF

Average VMAF

Average Quality	Ultrafast	Superfast	Veryfast	Faster	Fast	Medium	Slow	Slower	Veryslow	Placebo	Total Delta
Tears of Steel	89.20	92.00	93.29	95.45	95.59	96.22	96.43	96.56	96.67	96.65	8.38%
Sintel	88.29	92.66	93.85	95.84	95.99	96.38	96.56	96.68	96.83	96.75	9.68%
Big Buck Bunny	87.26	91.26	92.68	95.03	95.29	95.53	95.75	95.87	96.05	96.01	10.08%
Talking Head	95.19	92.55	93.66	94.90	94.86	95.18	95.29	95.43	95.51	95.39	3.20%
Freedom	91.95	91.15	92.63	94.58	94.51	95.37	95.59	95.84	96.15	96.04	5.48%
Haunted	91.30	88.61	89.43	91.30	91.08	91.98	92.08	92.35	92.49	92.45	4.38%
Screencam	90.92	92.56	93.52	94.75	94.75	94.70	94.77	94.86	94.92	94.91	4.41%
Tutorial	93.42	94.66	95.55	96.16	96.17	96.17	96.26	96.28	96.29	96.10	3.07%
Average	**90.53**	**91.37**	**92.59**	**94.52**	**94.56**	**95.11**	**95.28**	**95.46**	**95.62**	**95.55**	**6.08%**

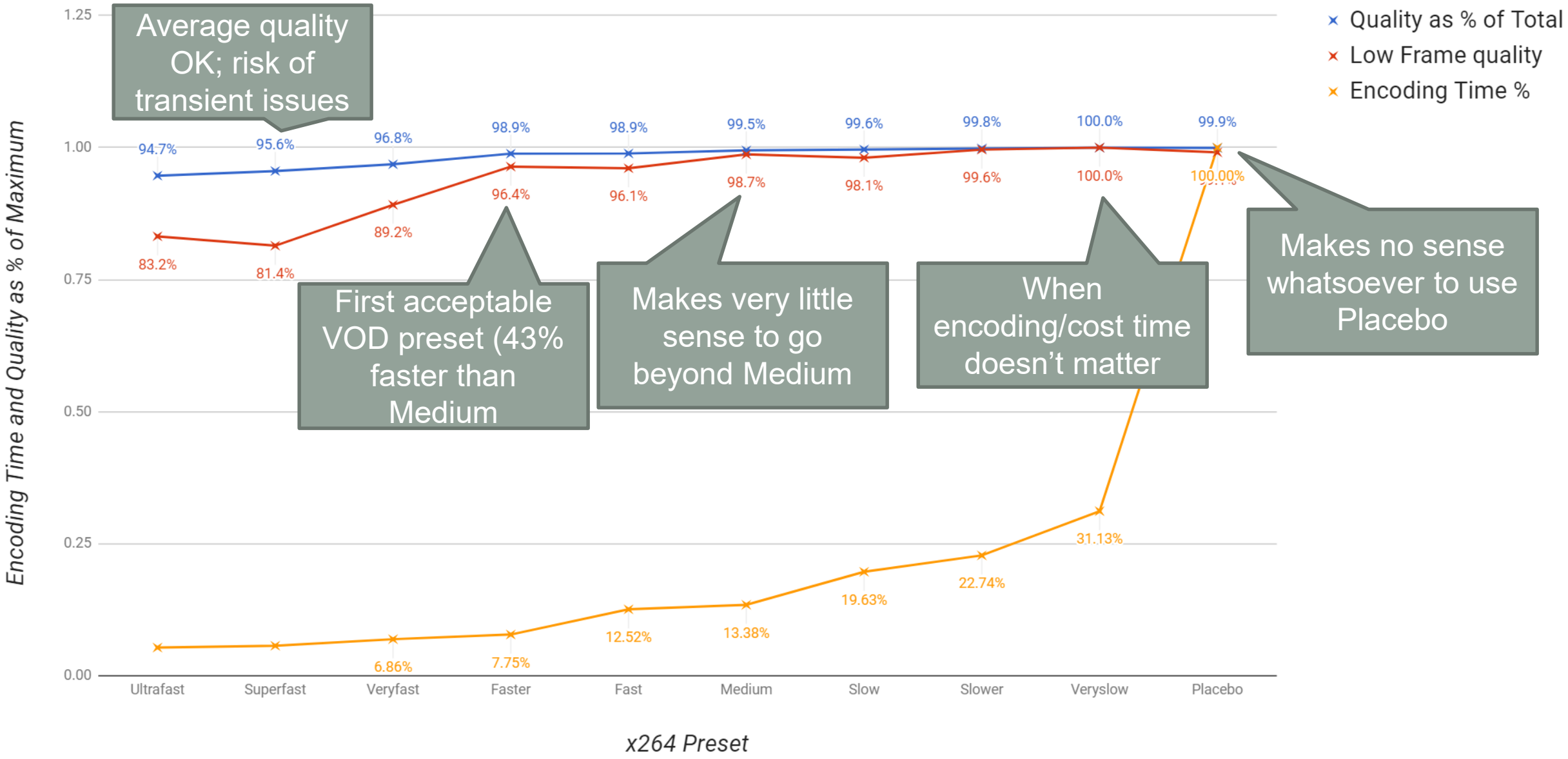
- Red is lowest quality
- Green highest quality
- Note top values – average 95.62 (not Placebo)
- Very slow averages best quality
  - But only 8% spread between best and worst

Low-Frame VMAF

Low Frame Quality	Ultrafast	Superfast	Veryfast	Faster	Fast	Medium	Slow	Slower	Veryslow	Placebo	Total Delta
Tears of Steel	70.16	74.82	77.67	84.51	85.02	85.34	85.44	86.38	85.33	85.10	23.12%
Sintel	68.77	69.79	74.93	79.12	80.41	82.27	81.90	82.98	84.89	82.61	23.45%
Big Buck Bunny	55.42	65.11	62.50	79.33	79.57	82.70	79.18	83.22	80.24	79.08	50.15%
Talking Head	88.90	61.43	88.53	91.62	91.32	92.11	92.03	92.49	92.16	91.37	50.56%
Freedom	76.49	82.79	83.96	87.59	87.29	88.72	89.00	89.35	90.28	90.05	18.03%
Haunted	60.36	57.18	62.69	64.62	61.63	67.33	67.74	68.64	72.08	72.28	26.42%
Screencam	56.16	68.53	71.00	76.39	77.44	77.06	78.04	79.26	78.04	75.21	41.12%
Tutorial	85.68	90.99	91.95	94.11	94.24	94.68	94.50	94.21	94.02	70.58	34.15%
Average	**70.02**	**68.52**	**75.05**	**81.13**	**80.88**	**83.08**	**82.55**	**83.84**	**84.16**	**83.41**	**33.37%**

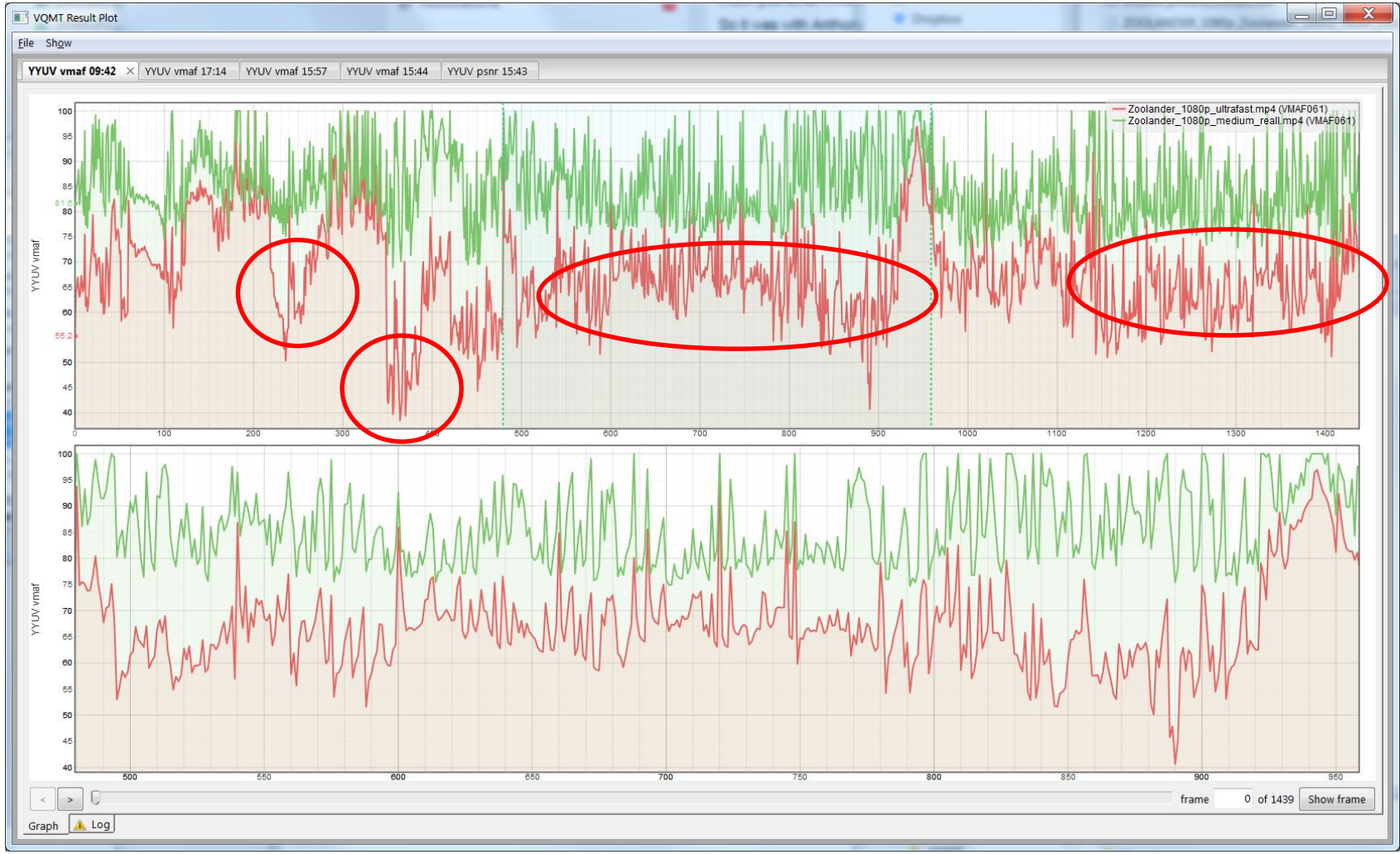
- Red is lowest quality
- Green highest quality
- Note top values – average 84.16 (not Placebo)
- Very slow averages best quality
  - 33% spread between best and worst

Average Quality, Low-Frame Quality and Encoding Time Per x264 Presets



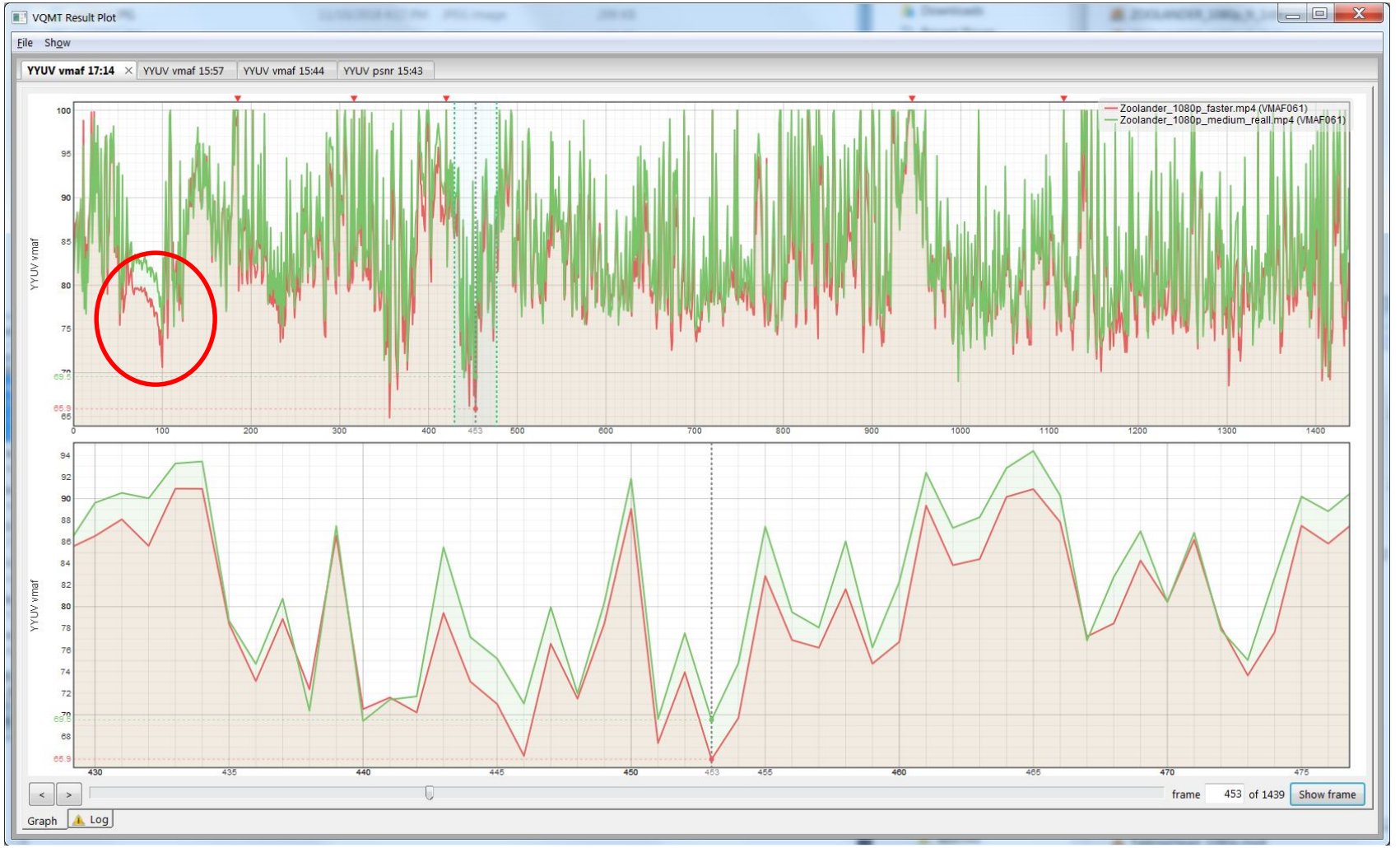
Check Results Plot – Ultrafast (red) vs Medium

- Multiple areas of significant differentiation
- Never use ultrafast (even in live)



Check Results Plot – Faster (red) vs Medium

- One problem area, but no major quality differences
- Fast should be acceptable starting point for VOD and live



Conclusions

- Faster is best preset for those seeking maximum throughput
- Makes very little sense to go beyond Medium when encoding cost/time is a concern
- Very slow delivers maximum average and low-frame quality; Placebo never seems to make sense

SSIMPLUS VOD Monitor Demos

- HEVC vs H.264
- Per-Title Encoding

HEVC vs. H.264

- Video reduced to save file size

Per-Title

- Video removed to reduce file size