BRIGHTCOVE

IMPROVING ABR ENCODING

by Adaptation to "True Resolution" of the Content

Yuriy Reznik, Karl O. Lillevold, Abhijith Jagannath, Nabajeet Barman*

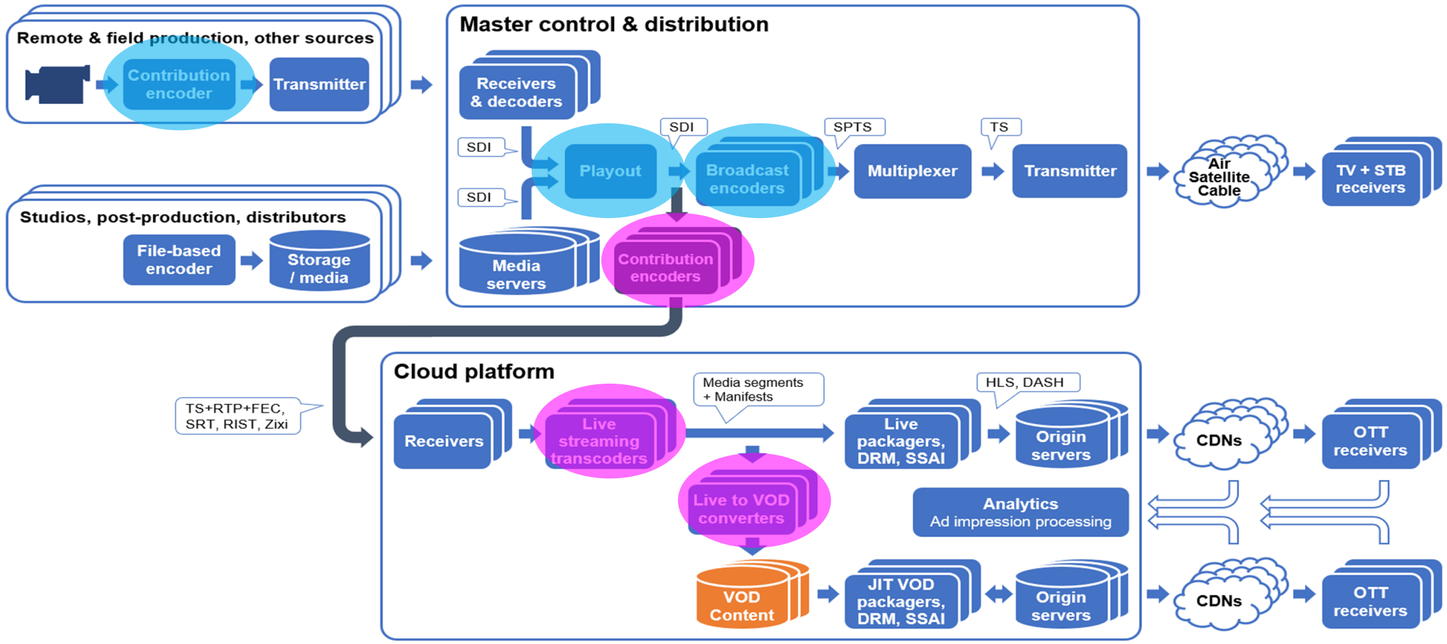
Brightcove, Inc, *Brightcove UK, Ltd

mhv/2023
ACM MILE-HIGH VIDEO
annual workshop

May 7-10, 2023, Denver, CO

THE PROBLEM

EXAMPLE HYBRID WORKFLOW



OBSERVATIONS

- Each stream becomes transcoded/converted several (1-4+) times
- Some conversions may result in **up-scaling**
- Using the upscaled content for distribution is suboptimal
  - More pixels and bits are sent out but quality does not improve!

EXAMPLES OF UP-CONVERSIONS

Formats	Width	Height	DAR	SAR	New target
SD/480i	352	480i	4:3	20:11	720p, 1080i, 1080p
	480	480i	4:3	4:3	
	528	480i	4:3	40:33	
	544*	480i	4:3	40:33	
	640	480i	4:3	1:1	
	704	480i	4:3	10:11	
	720*	480i	4:3	10:11	
	352	480i	16:9	80:33	
	480	480i	16:9	16:9	
	528	480i	16:9	160:99	
	544*	480i	16:9	160:99	
	640	480i	16:9	4:3	
	704	480i	16:9	40:33	
	720*	480i	16:9	40:33	
SD/576i	352	576i	4:3	24:11	
	480	576i	4:3	8:5	
	544*	576i	4:3	11:12	
	704	576i	4:3	12:11	
	720*	576i	4:3	12:11	
	352	576i	16:9	32:11	
	480	576i	16:9	32:15	
	544*	576i	16:9	64:33	
	704	576i	16:9	16:11	
	720*	576i	16:9	16:11	
HD/720p	960	720p	16:9	4:3	1080p, 4K
	1280	720p	16:9	1:1	
HD/1080i	1280	1080i	16:9	3:2	
	1440	1080i	16:9	4:3	
	1920	1080i	16:9	1:1	
HD/1080p	1920	1080	16:9	1:1	
Widescreen 1080p, 2K	1920	800p	2.4:1	1:1	1080p, 4K
	1920	816p	2.35:1	1:1	
	2048	858p	2.39:1	1:1	
	2048	864p	2.4:1	1:1	
	2048	1080p	1.9:1	1:1	

LOSS OF EFFICIENCY

THE CONTENT

CONVERSION

- Input is
  - 1920x800 widescreen (2.4:1) mezzanine
- Converted to
  - 1920x1080 HD video format (16:9 DAR)
- True resolution
  - 1422.2 x 800 pixels !

IMPACT ON STREAMING

1080P ENCODING LADDER

Rendition	Codec	Width	Height	True Width	True Height	Bitrate [kbps]	SSIM	Pr
1	h264	384	216	384	216	145	0.9421	0.0032
2	h264	512	288	512	288	267	0.9572	0.0039
3	h264	768	432	768	432	534	0.9656	0.0096
4	h264	1024	576	1024	576	1068	0.9745	0.0426
5	h264	1600	900	1422	800	2136	0.9773	0.1245
6	h264	1920	1080	1422	800	3763	0.9823	0.8084
Average				**1392**	**783**	**3386**	**0.9810**	
Storage						**7912**		

TRUE RESOLUTION-CONSTRAINED LADDER

Rendition	Codec	Width	Height	True Width	True Height	Bitrate [kbps]	SSIM	Pr
1	h264	384	216	384	216	145	0.9421	0.0030
2	h264	512	288	512	288	257	0.9556	0.0036
3	h264	768	432	768	432	488	0.9627	0.0082
4	h264	1024	576	1024	576	977	0.9720	0.0198
5	h264	1280	720	1280	720	1667	0.9769	0.0594
6	h264	1440	810	1422	800	2625	0.9815	0.8984
Average				**1394**	**784**	**2502**	**0.9806**	
Storage						**6159**		

IMPACT

- 26% higher use of bandwidth
- 22% higher use of storage

IMPACT ON ARTISTIC INTENT

ORIGINAL CONTENT

2X DOWNSCALED & UPSCALED

BRIGHTCOVE

EXISTING APPROACHES

CAE-TYPE SYSTEMS

- Per-title, content-aware, context-aware encoding techniques.
- Effective at reducing bandwidth, but keep declared resolution.
- Help somewhat, but not completely.

DYNAMIC RESOLUTION ENCODING

- Enabled in VVC as tool for more effective encoding
- Can also be implemented on a segment-level with HLS
- No always safe: may alter grain, fine details, may look inconsistent.
- Helps somewhat, but not completely.

ORIGINAL RESOLUTION DETECTORS

- Known since 1980s. Used in forensic analysis, restoration, etc.
- Many algorithms, but with limitations
  - Some only work for specific filters
  - Some only provide bounds for ranges of likely resolutions
  - Only few can detect true resolution after “super-resolution” upscalers
  - None is 100% reliable

REFERENCES

A. Aaron, et al., “Per-Title Encode Optimization,” December 2015. https://medium.com/netflix-techblog/per-title-encodeoptimization-7e99442b62a2

Y. Reznik, et al, "Optimal Design of Encoding Profiles for ABR Streaming," Packet Video Workshop, June 12 2018

C. Chen, et al, "Optimized Transcoding for Large Scale Adaptive Streaming Using Playback Statistics," ICIP, October 2018

Y. Reznik, et. al, "Optimizing Mass-Scale Multiscreen Video Delivery," SMPTE Motion Imaging Journal, 129 (3), 2020

ISO/IEC 23090-3:2022, "Information technology — Coded representation of immersive media — Part 3: Versatile video coding," ISO/IEC, 2022

X. Ducloux, et al., "Exploring the benefits of dynamic resolution encoding and support in DVB standards," IBC, September 2022

H. Farid, “Image forgery detection,” IEEE Signal processing magazine, 26(2), 2009

A. C. Gallagher, "Detection of linear and cubic interpolation in JPEG compressed images," 2nd Canadian Conference on Computer and Robot Vision, 2005

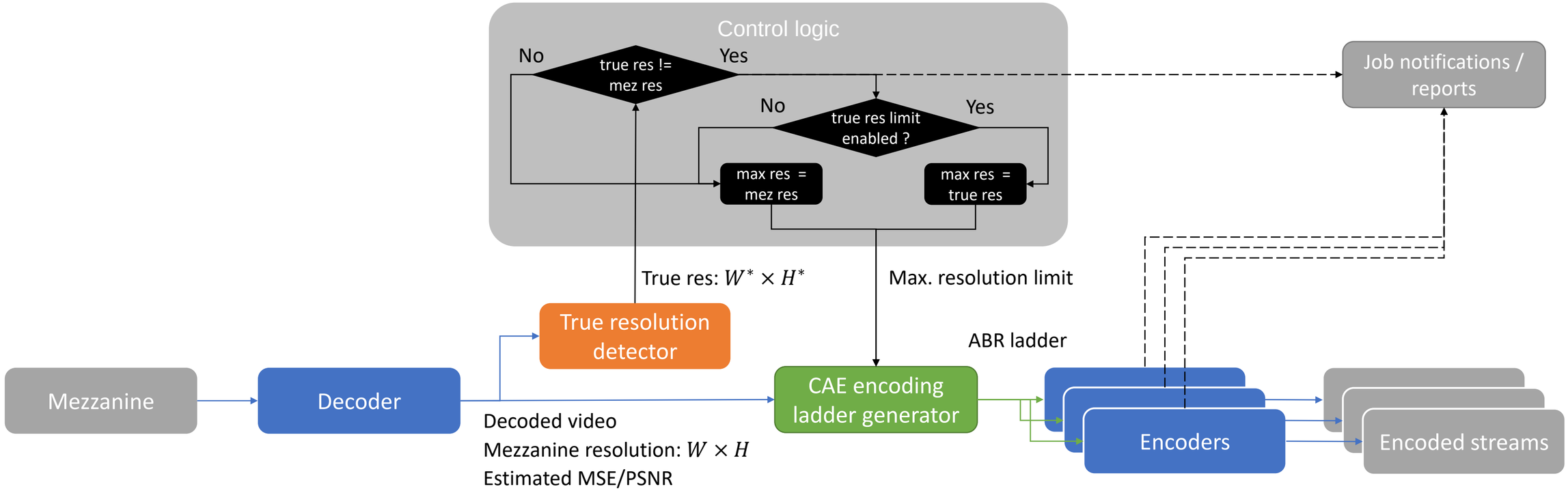
X. Feng, et al., "Normalized Energy Density-Based Forensic Detection of Resampled Image," IEEE Transactions on Multimedia, 14, (3), 2012

I. Katsavounidis, et al, "Native Resolution Detection of Video Sequences," SMPTE Technical Conference, 2015

Z. Yang, et al., "Native Resolution Detection for 4K-UHD Videos," IEEE Int. Symposium on Broadband Multimedia Systems and Broadcasting, 2020

PROPOSED SOLUTION

FLOW DIAGRAM



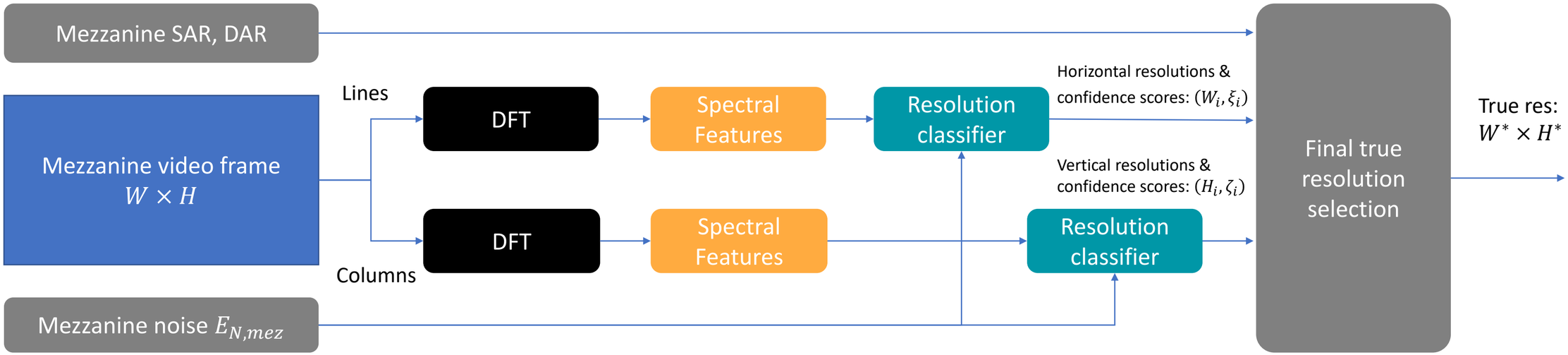
KEY FEATURES

- Uses new robust true resolution detector
- Uses context-aware encoding ladder generator (CAE)
- Delivers full range of savings: in bandwidth, storage, number of streams

BRIGHTCOVE

RESOLUTION DETECTOR

FLOW DIAGRAM

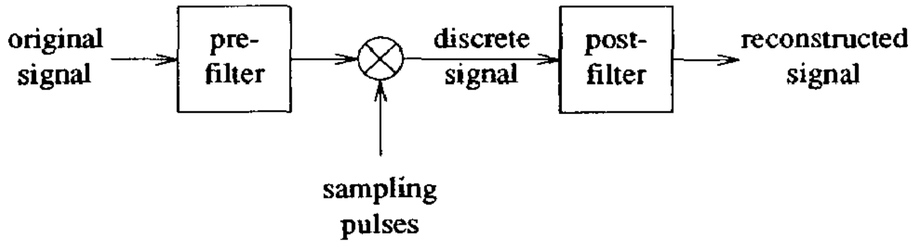


KEY FEATURES

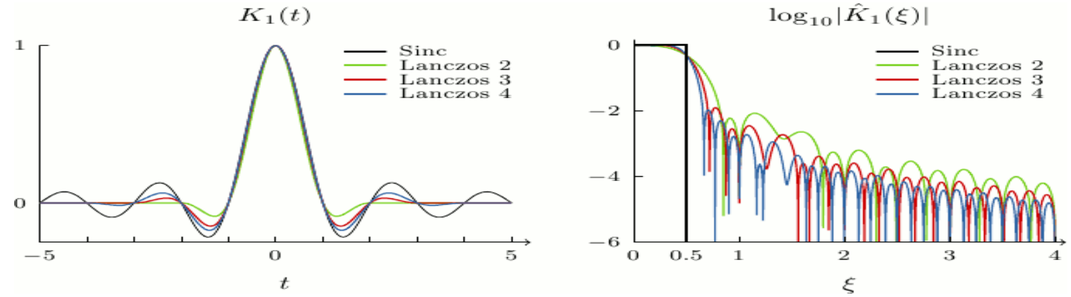
- Separate horizontal / vertical processing
- DFT-domain spectral features are used for detection
- Estimate of codec noise level in mezzanine is used as an extra parameter
- Horizontal and vertical detector produce initial lists of candidate resolutions in both directions
- Final selection block performs SAR/DAR consistency checks and selects best overall combination of horizonal and vertical resolution parameters

SPECTRAL FEATURES: BACKGROUND

SAMPLING & RECONSTUCTION

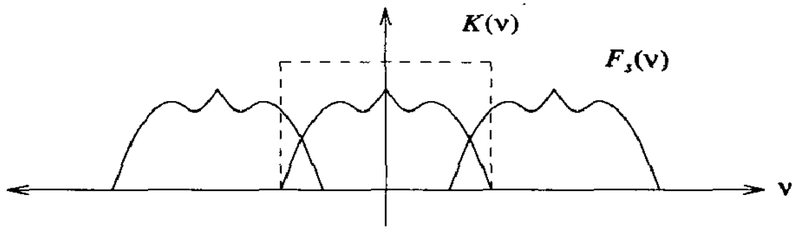


CLASSIC FILTERS

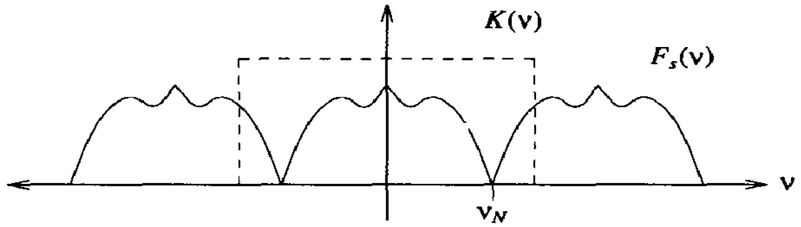


ALIASING ARTIFACTS

- Inadequate pre-filter (aliasing)

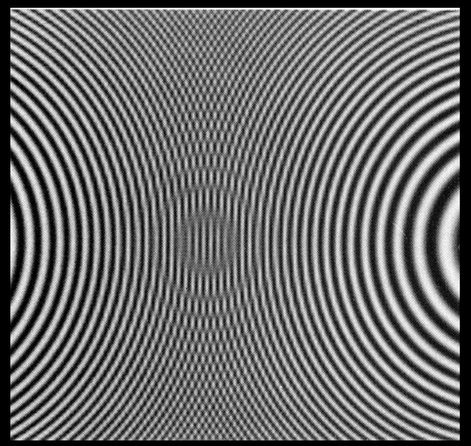


- Too broad post-filter (post-aliasing)

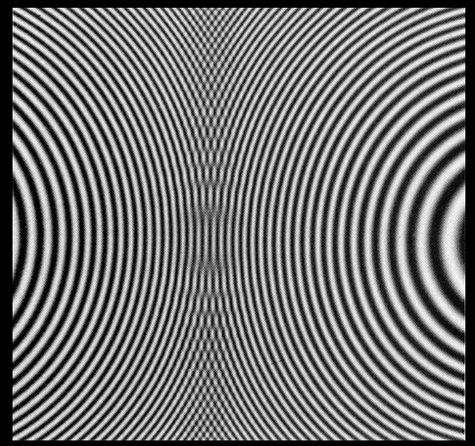


EXAMPLES

WITH ALIASING ARTIFACTS



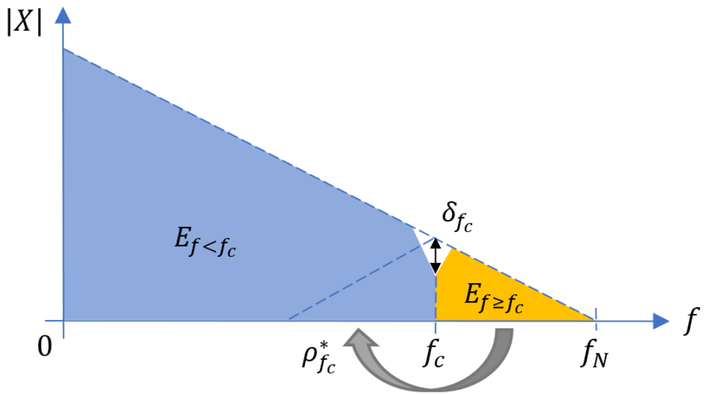
NEAR-IDEAL RECONSTUCTION



D. P. Mitchell and A. N. Netravali, "Reconstruction filters in computer graphics," Computer Graphics, vol. 22, no. 4, pp. 221–228, Aug. 1988

SPECTRAL FEATURES: DESIGN

CONCEPTUAL DIAGRAM



SAFETY CHECKS

- $E_{f\geq f_c} < \epsilon_0 + \epsilon_1 E_{N,Mez}$
- $E_{f\geq f_c}/E_{f<f_c} < \epsilon_2 + \epsilon_3 E_{N,Mez}$

NOTATIONS

- f_N – mezzanine Nyquist frequency
- f_c – candidate "true resolution" frequency under test
- $E_{f<f_c}, E_{f\geq f_c}$ – signal energies in bands before and after f_c
- $\rho^*_{f_c}$ – measure of presence of aliasing artifacts
- δ_{f_c} – measure of decline in the vicinity of f_c

FEATURES & PROCESS

- $\rho^*_{f_c}$ and δ_{f_c} are used to detect candidate "true resolution" frequencies f_c
- $E_{f<f_c}$, $E_{f\geq f_c}$ together with Mezzanine noise level $E_{N,Mez}$ are used to disqualify too aggressive choices of f_c

DETECTING ALIASING

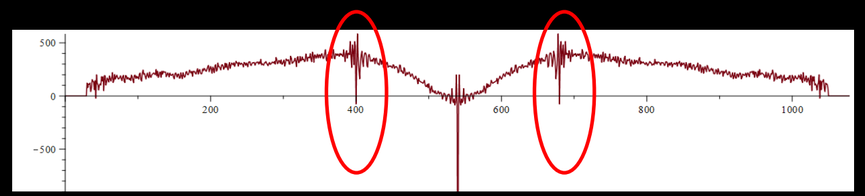
- Anti-correlation in folded imaginary spectrum:

$$\rho^*_{f_c} = \frac{\sum_u x_{im}[f_c - u] \cdot x_{im}[f_c + u]}{\sqrt{\sum_u x_{im}[f_c - u]^2 \cdot \sum_u x_{im}[f_c + u]^2}}$$

EXAMPLE

A FRAME WITH 800 LINES STRETCHED TO 1080

USING ANTI-CORRELATION METRIC:



BRIGHTCOVE

EXPERIMENT SETUP

TEST CORPUS

- 500 mezzanines (HD, SDR)
- 120 hours of programming
- 33 categories
- a sample of real-world content ingested in our OTT distribution workflow

ENCODED VARIANTS

- Encoded using standard HLS encoding ladder
- Encoded using Brightcove CAE
- Encoded using Brightcove CAE + True Resolution limit

METRICS COLLECTED

- Average bitrate
- Average storage
- Average number of resolutions
- Average true resolution delivered
- Average encoding quality delivered (SSIM)

EXAMPLE COMPARISON

REFERENCE HLS LADDER

Rendition	Codec	Profile	Width	Height	True Width	True Height	Bitrate [kbps]	SSIM	Pr
1	h264	High	416	234	416	234	145	0.9390	0.0049
2	h264	High	640	360	640	360	365	0.9604	0.0050
3	h264	High	768	432	768	432	730	0.9739	0.0076
4	h264	High	768	432	768	432	1100	0.9817	0.0336
5	h264	High	960	540	960	540	2000	0.9858	0.0710
6	h264	High	1280	720	1280	720	3000	0.9860	0.1261
7	h264	High	1280	720	1280	720	4500	0.9898	0.1298
8	h264	High	1920	1080	1422	800	6000	0.9874	0.1557
9	h264	High	1920	1080	1422	800	7800	0.9896	0.4586
Average					**1316**	**740**	**5705**	**0.9878**	
Storage							**25640**		

CAE LADDER

Rendition	Codec	Profile	Width	Height	True Width	True Height	Bitrate [kbps]	SSIM	Pr
1	h264	High	384	216	384	216	145	0.9421	0.0032
2	h264	High	512	288	512	288	267	0.9572	0.0039
3	h264	High	768	432	768	432	534	0.9656	0.0096
4	h264	High	1024	576	1024	576	1068	0.9745	0.0426
5	h264	High	1600	900	1422	800	2136	0.9773	0.1245
6	h264	High	1920	1080	1422	800	3763	0.9823	0.8084
Average					**1392**	**783**	**3386**	**0.9810**	
Storage							**7912**		

CAE + TR LADDER

Rendition	Codec	Profile	Width	Height	True Width	True Height	Bitrate [kbps]	SSIM	Pr
1	h264	High	384	216	384	216	145	0.9421	0.0030
2	h264	High	512	288	512	288	257	0.9556	0.0036
3	h264	High	768	432	768	432	488	0.9627	0.0082
4	h264	High	1024	576	1024	576	977	0.9720	0.0198
5	h264	High	1280	720	1280	720	1667	0.9769	0.0594
6	h264	High	1440	810	1422	800	2625	0.9815	0.8984
Average					**1394**	**784**	**2502**	**0.9806**	
Storage							**6159**		

BRIGHTCOVE

THE RESULTS

PER-CATEGORY COMPARISONS

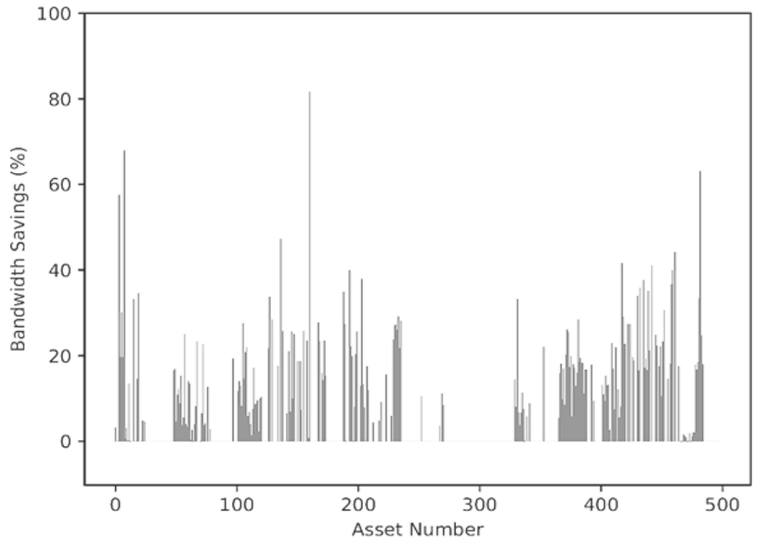
Content Category	Renditions			Storage [kbps]			Bandwidth [kbps]		
	REF	CAE	CAE+TR	REF	CAE	CAE+TR	REF	CAE	CAE+TR
Action	9.00	6.08	6.08	24361	8221	7890	5420	3477	3319
Adventure	9.00	6.17	6.17	25803	8964	8602	5741	3648	3516
Baseball	7.00	5.12	5.00	11477	4389	3631	3823	2086	1739
Basketball	8.61	6.06	5.76	23684	8036	5753	5530	3441	2513
Beach Volleyball	9.00	6.91	6.21	25858	12004	8432	5753	4506	3538
Boxing	9.00	6.00	6.00	25588	7590	6388	5693	3225	2622
Cartoon	9.00	5.84	5.70	25256	6690	5791	5619	2923	2545
Comedy	9.00	6.26	6.26	26655	9081	8700	5931	3516	3389
Cricket	7.00	5.00	4.32	12222	3231	2578	4072	1470	1296
Cycling	9.00	6.00	5.96	26272	8036	6805	6146	3289	2757
Documentary	9.00	6.50	6.47	24804	10226	9886	5519	3935	3787
Drama	9.00	6.17	6.17	26560	8492	8188	5910	3498	3362
Field Hockey	9.00	6.92	6.55	26180	11478	9808	5825	4242	3816
Football	7.45	4.67	4.67	14964	4055	3302	4346	1578	1375
Game Show	9.00	6.20	5.89	25137	8615	7870	5593	3566	3374
Gymnastics	9.00	6.00	6.00	24396	7111	6428	5707	2902	2902
Interview	7.07	4.40	4.05	11794	2591	2042	3839	1224	1003
Kids Channel	9.00	6.31	6.24	25292	9521	9141	5628	3829	3650
Late night show	9.00	6.28	5.50	24736	8722	6918	5786	3530	2923
Mixed Sports	9.00	6.84	6.74	24783	12260	11683	5514	4472	4328
News	9.00	6.51	6.23	26893	10038	9088	6291	3869	3666
Reality	9.00	6.50	6.43	25501	10044	9567	5674	3907	3760
Running	9.00	6.37	6.00	24663	9291	7352	5488	3828	3197
Scifi	9.00	6.18	6.12	24370	8933	8457	5422	3670	3519
Sitcom	9.00	5.99	5.98	24381	7284	6773	5425	3061	2863
Soap	9.00	6.14	6.03	25327	8185	7684	5635	3394	3239
Squash	9.00	6.00	6.00	25721	6990	6247	5723	3030	2711
Swimming	9.00	7.00	7.00	25823	13874	12993	5746	4784	4614
Tennis	7.00	5.00	5.00	11961	3553	3047	3984	1711	1450
Weightlifting	9.00	5.90	5.71	25915	6085	5165	5766	2616	2257
Overall	**8.67**	**6.04**	**5.87**	**23212**	**8119**	**7206**	**5418**	**3274**	**2967**
CAE vs Ref [%]		-30.30	-32.25		-65.02	-68.95		-39.57	-45.23
CAE+TR vs CAE [%]			-2.81			-11.25			-9.38

REF – encoded using reference HLS encoding profile; **CAE** – encoded using Brightcove CAE; **CAE+TR** – Brightcove CAE + true resolution detection

TEST CORPUS

- 500 mezzanines
- 120 hours of programming
- 33 categories
- real-world content

PER-ASSET SAVINGS



BRIGHTCOVE

CONCLUSIONS

TRUE RESOLUTION DETECTION WORKS

- Presented technique has been implemented in Brightcove CAE
- Customizable behavior. Safe to use.
- Delivers very appreciable bandwidth savings:
  - 9-11% savings on average
  - up to 30% on per-category basis
  - up to 80% on per-asset basis

BUT… THE RESULTS ARE ALSO ALARMING

- They show that a significant percentage of OTT content is up-scaled!
- In out data sets, we've observed:
  - up-converted 480i and 576i SD videos
  - content shot with 3:4 SAR pixels (e.g. 960x720 – possibly originating from old DVCPRO/HD cameras)
  - cropped and scaled wide-screen content (2.4:1 and 2.35:1 to 16:9 conversions)
  - up-converted 4K content, etc.
- As formats are evolving, we are likely to see even more up-converted content sent for distribution.
- Tools for detection and efficient encoding of such content may become progressively more important.

BRIGHTCOVE®

THANK YOU