CREATING THE LIVING NETWORK™

JVET-E0071
AhG8: Viewport-based subjective evaluation of 360-degree video coding

Philippe Hanhart, Yuwen He, Yan Ye (InterDigital)

© 2016 InterDigital, Inc. All Rights Reserved.

INTERDIGITAL®

Introduction

- Experimental results demonstrate average gain of 23-24% for JEM-4.1 over HM-16.14 for ERP format based on various objective metrics (JVET-D1030)
- Correlation between objective quality metrics and perceived visual quality is still an open question for 360-degree video
  - Subjective quality assessment remains the ultimate means to assess visual quality
- Subjective quality evaluations can be conducted using HMD devices or regular 2D displays using viewport rendering
  - HMD is VR dedicated device for VR, whereas usage on 2D displays expected to continue
- This contribution reports results of two subjective evaluations comparing viewports rendered from 360-degree video sequences encoded with HM-16.14 and JEM-4.1 in ERP format

© 2016 InterDigital, Inc. All Rights Reserved.

Subjective evaluation

- Goal: evaluate visual quality improvements for JEM-4.1 over HM-16.14 in Random Access configuration following 360 CTC (JVET-D1030)
- Task: design suitable test method for fair comparison
  - Viewport based evaluation on 2D displays
- Subjective experiments
  - Experiment #1: matching QP (QP=37) – 5 expert viewers (preliminary experiment)
  - Experiment #2: matching bit rate (JEM @ QP=37) – 9 expert viewers
- Test environment
  - Dark room with grey walls and black curtains
  - Calibrated Samsung UN55HU6950 55” 4K LCD TV
  - 1.6H viewing distance

© 2016 InterDigital, Inc. All Rights Reserved.

Test method

- Stimulus comparison method
- Side-by-side comparison
  - HM vs JEM (random position)
  - Minimize visual working memory limitations
- Two presentations of each pair
  - Increased sensitivity
- Semantic adjectival categorical judgement scale

Left much better	Left better	Left slightly better	Same	Right slightly better	Right better	Right much better
☐	☐	☐	☐	☐	☐	☐

© 2016 InterDigital, Inc. All Rights Reserved.

Dataset – Experiment #1: matching QP

Sequence name	QP		bit rate [kbps]			luma PSNR [dB]			
						VP1		VP2	
	HM	JEM	HM	JEM	JEM-rate / HM-rate (%)	HM	JEM	HM	JEM
Train_le	37	37	1519.4	1291.9	85.0	33.49	34.20	33.02	33.69
SkateboardingTrick_le	37	37	2009.6	1769.9	88.1	30.16	30.62	38.47	39.08
SkateboardInLot	37	37	1865.0	1598.9	85.7	32.84	33.41	34.10	34.69
ChairLift	37	37	1508.5	1295.2	85.9	36.43	37.00	33.90	34.46
KiteFlite	37	37	2676.6	2506.7	93.7	34.37	34.86	32.93	33.38
Harbor	37	37	910.4	854.3	93.8	34.33	34.79	36.17	36.79
PoleVault_le	37	37	1649.8	1512.4	91.7	33.89	34.33	32.81	33.24
AerialCity	37	37	822.0	709.8	86.4	34.57	35.19	33.91	34.53
DrivingInCity	37	37	1118.7	907.3	81.1	33.34	34.08	33.28	33.90
DrivingInCountry	37	37	2129.8	1806.8	84.8	30.73	31.48	29.88	30.47

© 2016 InterDigital, Inc. All Rights Reserved.

Dataset – Experiment #2: matching bit rate

Sequence name	QP		bit rate [kbps]			luma PSNR [dB]			
						VP1		VP2	
	HM	JEM	HM	JEM	JEM-rate / HM-rate (%)	HM	JEM	HM	JEM
Train_le	38	37	1331.8	1291.9	97.0	32.90	34.20	32.45	33.69
SkateboardingTrick_le	37.5	37	1865.6	1769.9	94.9	29.94	30.62	38.24	39.08
SkateboardInLot	38	37	1587.4	1598.9	100.7	32.39	33.41	33.73	34.69
ChairLift	38	37	1275.2	1295.2	101.6	36.02	37.00	33.53	34.46
KiteFlite	37.5	37	2500.4	2506.7	100.3	34.12	34.86	32.66	33.38
Harbor	37.5	37	847.7	854.3	100.8	34.11	34.79	35.92	36.79
PoleVault_le	37.5	37	1551.9	1512.4	97.5	33.63	34.33	32.55	33.24
AerialCity	38	37	733.8	709.8	96.7	34.08	35.19	33.40	34.53
DrivingInCity	38.5	37	914.6	907.3	99.2	32.56	34.08	32.58	33.90
DrivingInCountry	38	37	1786.6	1806.8	101.1	30.22	31.48	29.44	30.47

* HM was re-encoded at roughly the same bit rate as JEM

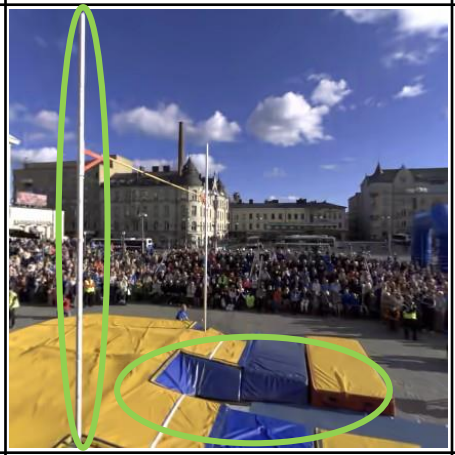
© 2016 InterDigital, Inc. All Rights Reserved.

	Train	SkateboardingTrick	SkateboardInLot	ChairLift
VP1				
VP2				

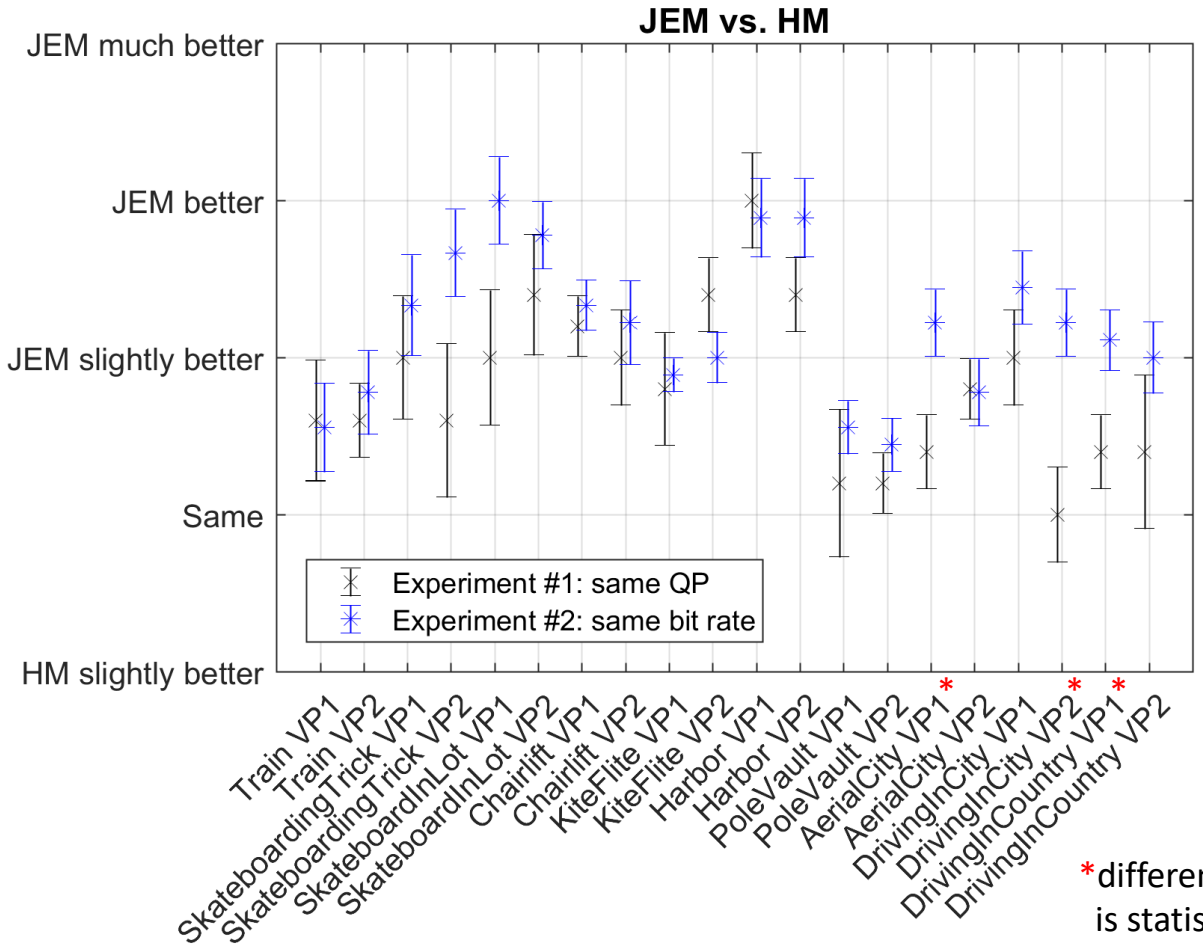
© 2016 InterDigital, Inc. All Rights Reserved.

	KiteFlite	Harbor
VP1		
VP2		

© 2016 InterDigital, Inc. All Rights Reserved.

	PoleVault	AerialCity	DrivingInCity	DrivingInCountry
VP1				
VP2				

© 2016 InterDigital, Inc. All Rights Reserved.



*difference between exp #1 and exp #2 is statistically significant ($p \leq 0.05$)

© 2016 InterDigital, Inc. All Rights Reserved.

Observations – sequences

- Improvements are most visible on Harbor and SkateboardInLot
- Improvements are less visible for 4K ERP because of viewport size (856x856 for 4K ERP vs. 1816x1816 for 8K ERP)
- AerialCity, PoleVault: few details as objects are relatively far from camera
- DrivingInCountry: lot of camera motion probably due to lack of camera stabilization. Viewport looking at river is rather unpleasant to watch; tends to create some sort of simulator sickness
- Sequences with fast motion (Train, SkateboardInLot): quality difference is difficult to notice during playback because of temporal masking. Improvements can be perceived when displayed as still frames

© 2016 InterDigital, Inc. All Rights Reserved.

Observations – general

- Stitching artefacts can be disturbing, especially for non-expert viewers or viewers not familiar with the content, and may influence judgement
  - Recommendation #1: select viewports where stitching artefacts are minimized
  - Recommendation #2: avoid single stimulus test methods
- Rendered viewports with FOV=75° appear a little blurred in the corners
  - Recommendation: larger FOV values should not be used to avoid over-blurring

© 2016 InterDigital, Inc. All Rights Reserved.

Dynamic viewports

- Good alternative for rather static content
- For sequences with relatively fast motion, might introduce some visual discomfort or create simulator sickness if not selected carefully
- Should mimic natural viewer behavior (i.e., more horizontal than vertical viewport movement)
- Dynamic viewports can change relative motion (ex: Train)
  - Considering temporal masking, visibility of video compression artefacts might be affected by change in relative motion
  - Recommendation: dynamic viewports should be moving in similar direction as main object, rather than against it

© 2016 InterDigital, Inc. All Rights Reserved.

Bullet time viewports

- Observation: for sequences with fast motion, compression artefacts and improvements from JEM over HM were more visible on still frames than during playback due to temporal masking
- Proposal: consider dynamic viewport defined at a fixed time instance
  - Resulting effect similar to ‘bullet time’ visual effect used in “The Matrix”
  - Similar idea previously used for quality evaluation for free-viewpoint TV
- Mimics user behavior of exploring content while video is paused
  - Real life application of 360-degree video
- Patch for 360Lib is provided to enable bullet time viewport rendering
  - ViewPortFile: specify viewport parameters
  - ViewPortType: new parameter

© 2016 InterDigital, Inc. All Rights Reserved.

Viewports comparison

- Duration and framerate:
  - Static/dynamic viewports: same as source 360° video
  - Bullet time viewports: can be set **independently** from source 360° video to achieve smooth and pleasant rendering
- Static viewports: selected viewport location influences results
- Bullet time viewports: selected time instance influences results

Viewport type	Space coverage	Time coverage	Visual comfort	Comment
Static	Poor	Good	Good	Suitable for any type of content
Dynamic	Average – Good	Good	Poor – Good	Not recommended for sequences with fast motion Visual comfort is somewhat inversely proportional to space coverage
Bullet time	Average – Good	Poor	Good	Suitable for any type of content

© 2016 InterDigital, Inc. All Rights Reserved.

Conclusion

- Subjective evaluation results demonstrates that JEM can achieve visible quality improvements over HM on all sequences
  - Areas where visual difference is observed are reported for each content
- We propose to add bullet time viewports to 360Lib software (patch attached)
  - We would like to request a viewing session to demonstrate some bullet time viewports in comparison with static and dynamic viewports
- Assessing the whole 360-degree video requires many viewports, thus we suggest to consider different viewport types (static, dynamic, and bullet time) with different characteristics for each content

© 2016 InterDigital, Inc. All Rights Reserved.

THANK YOU!

© 2016 InterDigital, Inc. All Rights Reserved.

INTERDIGITAL