



JCTVC-F158: Resolution switching for coding efficiency and resilience

Thomas Davies

Resolution switching in video conferencing

- Resolution switching is very commonly used in video conferencing
- Rate adaption due to changing network conditions (bandwidth and errors)
 - E.g. based on feedback from the receiver or intermediate node
- Composition (picture in picture for non-speakers)
 - Speakers are switched to full-frame
 - Don't want non-speakers transmitting at full bandwidth
- Video conferencing equipment has scaling infrastructure for handling these applications

Problem statement

- The AVC method of changing resolution requires an IDR frame
- Intra frames of the same quality as predicted frames are many times larger ($>10x$)
 - This ratio has increased further in HEVC
- Therefore we have to choose between
 1. A good quality IDR frame imposing large latencies and extra complexity, which may make things worse, or
 2. A very poor quality IDR showing a visible glitch, or
 3. Not changing resolution

Proposal and tool description

- We propose H263-style reference picture re-scaling
Currently just 2:1 and 1:2
- Signal a Picture Resolution Index (PRI) in the PPS
- When a reference picture has a different PRI to the current picture, re-scale the reference
Turn off temporal MV prediction in this case
- Use simple separable FIRs for up- and down-scaling
More complex filters could be used for display
- BD-Rate Gains of 5.4-7.4% for Low Delay common conditions, with 15.6-21.6% gain in Class E, are reported

Filtering

- Downconversion filter (luma and chroma):
 $(-1, 9, 16, 9, -1)/32$
- Upconversion filter:
 $(-3, 0, 19, 32, 19, 0, -3)/32$

Experimental set-up

- Modified HM3.1 to support PRI and multiple resolutions
- Resolution is switched down at frame 20. Previous 4 frames are marked as non-reference (assumed lost)
- Compare Low Delay configurations with:
 - A: inserted IDR frame at frame 20, no prediction from earlier frames, GOP structure re-started
 - B: GOP structure continued, with reference re-scaling
- BD-rate comparison performed **on the low-res segment**
 - PSNR and bit rate measured at low-res level
- Comparisons done with various QP settings for IDR frames: QP+0, QP+6, QP+12
 - QP+12 gives most realistic Intra frame size

Class E intra frame size ratios

- For IDR $QP=QP+0$, the ratios of low res Intra frame size to mean frame size are:

	Vidyo1	Vidyo3	Vidyo4	Mean (frame periods)	Milliseconds
QP=22	14.9	13.1	13.9	14.0	248
QP=37	26.9	24.7	25.1	25.7	428

- For IDR $QP=QP+12$, the ratios of low res Intra frame size to mean frame size are:

	Vidyo1	Vidyo3	Vidyo4	Mean (frame periods)	Milliseconds
QP=22	4.1	3.5	3.5	3.7	62
QP=37	5.2	4.2	5.0	4.8	80

Results : large I frames

	Low delay B HE			Low delay B LC		
	Y	U	V	Y	U	V
Class A						
Class B	-5.2	-10.2	-11.2	-5.1	-9.9	-10.5
Class C	-2.2	-4.2	-4.1	-2.1	-4.3	-3.7
Class D	-1.7	-5.7	-4.3	-1.6	-2.0	-2.8
Class E	-16.4	-27.7	-33.3	-15.6	-31.1	-36.6
Overall	-5.7	-10.8	-11.8	-5.4	-10.5	-11.8
Enc Time[%]	100%			100%		
Dec Time[%]	92%			92%		

	Low delay P HE			Low delay P LC		
	Y	U	V	Y	U	V
Class A						
Class B	-4.9	-9.8	-11.4	-5.1	-10.2	-11.3
Class C	-2.1	-4.4	-3.9	-2.2	-5.1	-4.6
Class D	-1.6	-7.2	-5.4	-1.7	-3.0	-3.1
Class E	-16.5	-29.0	-35.3	-16.3	-32.8	-38.2
Overall	-5.6	-11.4	-12.5	-5.6	-11.4	-12.6
Enc Time[%]	100%			100%		
Dec Time[%]	92%			93%		

Results – I frames QP+12

	Low delay B HE			Low delay B LC		
	Y	U	V	Y	U	V
Class A						
Class B	-6.6	-12.3	-12.8	-6.7	-12.0	-12.0
Class C	-2.9	-6.0	-5.6	-2.8	-5.8	-5.2
Class D	-2.2	-7.0	-6.3	-2.3	-4.4	-5.0
Class E	-21.5	-30.5	-37.3	-20.9	-33.6	-39.9
Overall	-7.4	-12.8	-14.0	-7.3	-12.6	-13.8
Enc Time[%]	100%			100%		
Dec Time[%]	98%			98%		

	Low delay P HE			Low delay P LC		
	Y	U	V	Y	U	V
Class A						
Class B	-6.4	-12.3	-13.0	-6.6	-12.2	-12.4
Class C	-2.7	-5.7	-5.4	-2.8	-6.4	-5.4
Class D	-2.3	-5.8	-7.4	-2.3	-4.1	-4.8
Class E	-21.5	-32.4	-40.2	-21.6	-35.1	-42.0
Overall	-7.3	-12.8	-14.8	-7.4	-13.0	-14.3
Enc Time[%]	100%			100%		
Dec Time[%]	93%			96%		

Frame 20 QP37 – inserted IDR (QP+12)



Frame 20 QP37 – predicted



Conclusions

- Resolution switching is widely employed for maximising quality in varying network conditions
- AVC is restrictive in requiring an IDR frame when this happens. Often a usable LTR exists.
- Intra frames are expensive and very problematic for low-delay applications. In HEVC this will be worse.
- We would like to propose that cross-resolution prediction is integrated into HEVC SW and WD to support these use-cases.

