A Scalable Video Coding Dataset and Toolchain for Dynamic Adaptive Streaming over HTTP

Christian Kreuzberger, Daniel Posch and Hermann Hellwagner Institute of Information Technology (ITEC) Alpen-Adria-Universität (AAU) Klagenfurt, Austria

Motivation

- Inherent caching on network nodes
- Dash propose to use many representations at various quality
- Roughly 50% of storage space could be saved

Scalable video Coding(SVC)

- Videos are split into several independent layers
- Base layer (BL)、 Enhancement layer (EL)
- Provide EL layers only when resource are available

TOOLCHAIN

- Proprietary MainConcept SVC/H.264 Encode in favour of the JSVM reference encoding software

Encoder Settings

- Force I-Frames to be at the beginning of every segment
- Segments independently decodable
- Video resolution, bitrate and QP need to be chosen per layer

De-Multiplexing the H.264/SVC Bitstream

- Access Units (AUs)
- AU represents a frame and Network Abstraction Layer Units (NALUs)
- NALU contains header information and describes the frame for a specific layer

De-Multiplexing the H.264/SVC Bitstream

Type	(D,T,Q)	Frame	
H.264/SVC Header		(init-file)	
New AU			
AVC-I	(0,0,0)	1	→ BL
SVC-I	(1,0,0)	1	→ EL 1
SVC-I	(2,0,0)	1	→ EL 2
New AU			
AVC-B	(0,1,0)	2	→ BL
SVC-B	(1,1,0)	2	→ EL 1
SVC-B	(2,1,0)	2	→ EL 2
New AU			
AVC-B	(0,0,0)	3	→ BL
SVC-B	(1,0,0)	3	→ EL 1
SVC-B	(2,0,0)	3	→ EL 2
⋮	⋮	⋮	
New AU			
AVC-P	(0,1,0)	48	→ BL
SVC-P	(1,1,0)	48	→ EL 1
SVC-P	(2,1,0)	48	→ EL 2
H.264/SVC Header		(segment border)	
New AU			
AVC-I	(0,0,0)	49	→ BL
SVC-I	(1,0,0)	49	→ EL 1
SVC-I	(2,0,0)	49	→ EL 2
⋮	⋮	⋮	
End of Stream			

Bitrate and Spatial Resolution

- Using too many enhancement layers would have a negative impact on the client

- Encode the content with MainConcept's VBR with four different variants

LayerId	Resolution	AVC bitrate	SVC bitrate
I.1.BL	640x360	600 kbps	600 kbps
I.1.EL1	640x360	900 kbps	990 kbps
I.1.EL2	640x360	1250 kbps	1500 kbps
I.1.EL3	640x360	1600 kbps	2075 kbps
I.2.BL	1280x720	1500 kbps	1500 kbps
I.2.EL1	1280x720	2500 kbps	2750 kbps
I.2.EL2	1280x720	4000 kbps	4800 kbps
I.2.EL3	1280x720	6000 kbps	7800 kbps
I.3.BL	1920x1080	4000 kbps	4000 kbps
I.3.EL1	1920x1080	5000 kbps	5500 kbps
I.3.EL2	1920x1080	6000 kbps	7200 kbps
I.3.EL3	1920x1080	8000 kbps	10400 kbps

Table 3: Variant I – Resolution and bitrates based on [7], with 10 % overhead per layer [7]

Resolution	AVC bitrates [kbps]	SVC bitrates [kbps]
480x360	180, 220, 370	180, 242, 444
1280x720	780, 1000, 1500	780, 1100, 1800
1920x1080	2000, 2900, 3190	2000, 3190, 5040

Table 4: Variant II – Resolution and bitrates based on [10], with 10 % overhead per layer [7]

DATASET ANALYSIS

- The SVC encoded videos were decoded and analyzed in terms of PSNR and SSIM

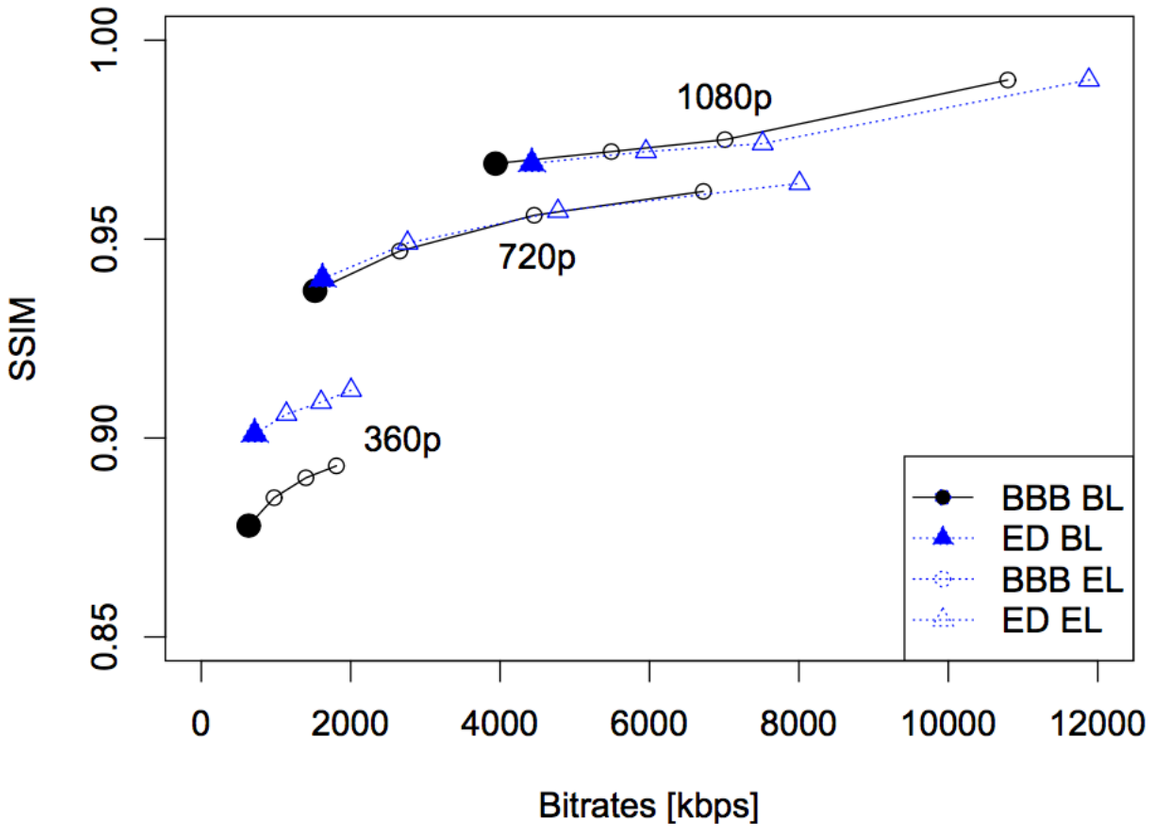

Figure 2: SSIM values (at 1080p) for BBB and ED (SVC Variant I)

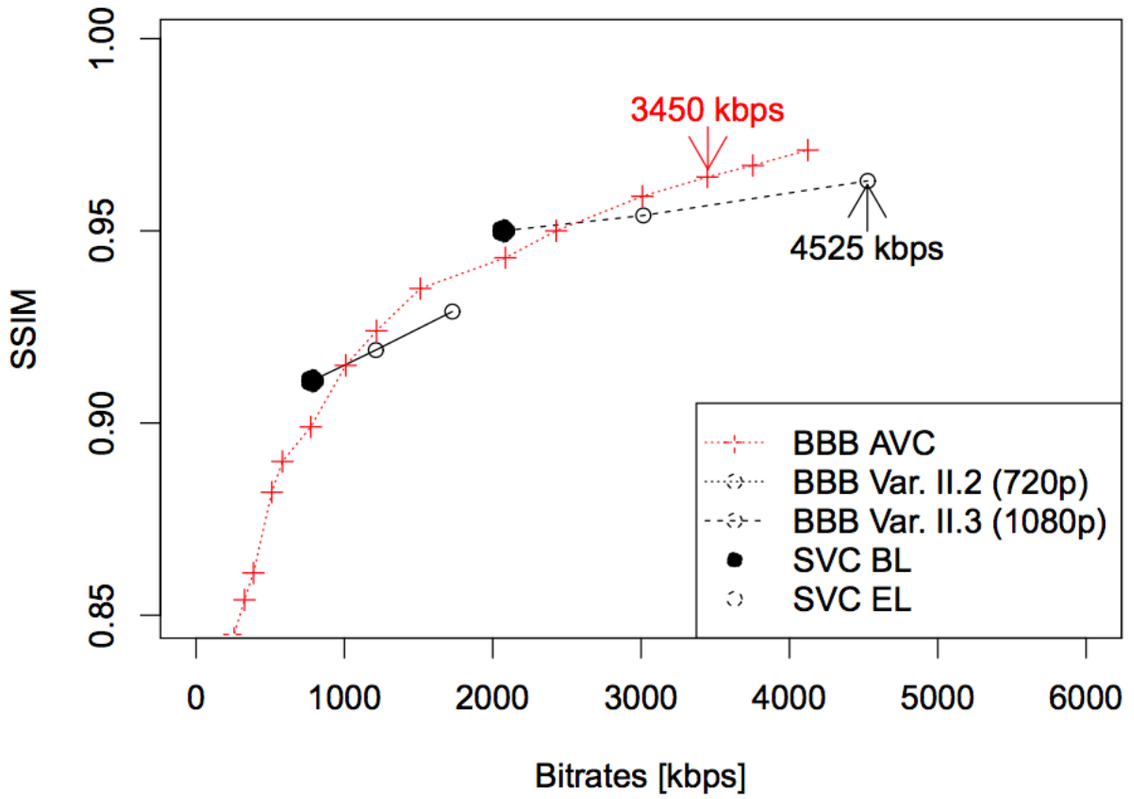

Figure 3: SSIM values (at 1080p) for BBB for SVC Variant II and DASH with AVC