An adaptive video streaming framework for Scalable HEVC (SHVC) standard

Sarat Rakngan, Thanuja Mallikarachchi, and Anil Fernando
Centre for Vision Speech and Signal Processing, University of Surrey, United Kingdom
Email: {sr00636, d.mallikarachchi, w.fernando}@surrey.ac.uk

***Abstract*—This paper presents an implementation of a Media Aware Network Element (MANE) for dynamic video content adaptation in Scalable HEVC (SHVC) video streaming. The experimental results discuss the varying quality-to-playback time ratio and decoding power consumption with random access period in SHVC encoding under fluctuating and persistent network bandwidth conditions.**

I. Introduction

The increasing video resolutions and exorbitant mobile video consumption (expected to reach over 82% of the overall consumer Internet traffic by 2021 [1]) impose several strains on the constrained network bandwidth. In this regard, the state-of-the-art High Efficiency Video Coding (HEVC) video coding standard, with its superior coding performance over its predecessor (H.264/AVC) [2], is expected to reduce the ever increasing bandwidth demands. Yet, the media content delivery under fluctuating network conditions to heterogeneous consumer electronic devices is a compelling challenge for the content creators as well as distributors.

In this context, adaptive streaming with scalable bit streams is seen as a potential solution, where the video contents are encoded into multiple layers; a Base Layer (BL) and one more Enhancement Layer (EL). The recently introduced Scalable HEVC (SHVC) extension [3] surpasses the shortcomings of its predecessor (SVC) [4][3] and therefore, is expected to be utilized within the existing media processing workflows.

Numerous researches exist in the literature that investigate the adaptive streaming techniques using SHVC. For example Hamidouche *et al*. [5] and Hadhrami *et al*. [6] propose a video streaming framework and an evaluation toolset for SHVC, respectively. A more comprehensive SHVC based video stream adaptation in wireless networks is proposed in [7], which uses a Media Aware Network Element (MANE) to extract layer information and slice the SHVC bit streams depending on the network bandwidth. However, a significant scope still exists to further investigate the variation of video quality and decoding power consumption of SHVC bit streams, which are encoded with different random access periods, under fluctuating and persistent bandwidth conditions.

The remainder of this paper is organized as follows. Sec. II and Sec. III describe the proposed streaming framework and experimental results, respectively. Finally, Sec. IV presents conclusions and potential future work.

This work was supported by the CONTENT4ALL project, which is funded under European Commission's H2020 Framework Program (Grant number: 762021).

II. Proposed Streaming Framework

The proposed adaptive streaming framework constitutes a media server, MANE, and a client device (Ref. Fig. 1). The media server is used to encode the raw video sequences. The client device receives the SHVC streams through MANE and an *openHEVC* [8] based player is configured to listen to a TCP/IP port to receive, decode and play the video data. MANE is designed to perform three key functions; bandwidth monitoring, NAL extraction and layer adaptation.

A. Bandwidth monitoring

MANE and client device are installed with *iperf* [9] utility which is executed as a parallel thread to monitor the network bandwidth of the wireless connection. The bandwidth measured is recorded in a cyclic buffer and the current bandwidth is estimated as a moving average using,

$$\mathrm{BW}_{est} = \frac{1}{\widetilde{W}} \sum_{i=0}^{n-1} \mathrm{BW}(t) H(t-n), \tag{1}$$

where the windowing function is given by,

$$H(t) = \begin{cases} 1 + \frac{t}{W}, & 0 \leq t \leq -W \\ 0, & \text{otherwise} \end{cases} . \tag{2}$$

Here, $\widetilde{W}$ is the area under the curve $H(t)$ and W is the window size, which is maintained at $W = 10$.

B. NAL extraction and layer adaptation

The encoded SHVC bit streams are encapsulated in Network Adaptation Layer (NAL) units which constitute a header and raw data payload. NAL units are first analyzed through the MANE's NAL parser which extracts the NAL unit type (i.e., random access frame, end of access unit etc.) and layer identification (ID). NAL unit type and layer ID are then processed by the layer extraction module which determines whether the NAL units should be sent to the client via the TCP/IP connection. For example, if the current bandwidth is less than EL bit rate, then the layer extraction module is instructed to transmit only the NAL units corresponding to the BL. The NAL unit type is compared before each layer switch to determine the beginning of a random access frame, and to skip the corresponding non-decodable leading pictures in order to avoid decoding errors at the client device.

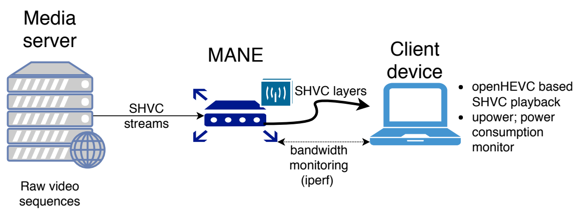

Fig. 1. Experimental setup consisting of media server, MANE and client device.

TABLE I
TEST SEQUENCES AND ENCODING BIT RATES

class	Sequence	BL bit rate	EL bit rate
A	People On Street	5 Mbps	10 Mbps
B	Kimono	2 Mbps	4 Mbps
C	Race horses	625 kbps	1.2 Mbps
D	Blowing bubbles	250 kbps	500 kbps
E	Four people	1.25 Mbps	2.5 Mbps

III. EXPERIMENTAL RESULTS AND DISCUSSION

Fig. 1 illustrates the simulation environment. In this case, MANE and client device are both simulated using two machines running Ubuntu 16.05 operating system. Table III depicts the BL and EL bit rates used to encode the test video sequences. Here, BL raw video sequences are formulated with half the resolution of the EL raw sequences. The test sequences are encoded with SHM 4.1 [10] reference encoder using *encoder lowdelay P main.cfg* with intra-periods configured at 8 and 32, using Clean Random Access (CRA) frames as random access points. A fluctuating network environment is simulated between the two TCP/IP network ports between MANE and client (Fig. 1) using Linux traffic control utilities [11]. The network bandwidth is then randomly varied at a frequency of every 5 and 20 seconds (f_{bw}) between base ($1.1 \times \mathrm{BW}_{bl}$) and enhancement ($1.1 \times \mathrm{BW}_{el}$) layer bit rates.

The quality of the received video streams are measured using Peak Signal-to-Noise Ratio (PSNR). The power consumption of the client device is measured in watt-hours (wh) using *upower* [12] Linux utility. The playback times for the sequences differ with the variation of the bandwidth and layers selected by the MANE. Hence, two metrics are defined as Quality:Time ratio and energy consumption rate ratio given by QL_t(dB/s) and PW_t(mAh/s), respectively.

It can be observed that when the bandwidth fluctuates at a higher frequency ($f_{bw} = 5$), the QL_t ratio is high when the intra-period is set to 8. This encoding configuration allows the SHVC bit streams to accumulate an increased number of random access points, allowing the proposed MANE to rapidly switch between layers aligning with the actual bandwidth changes. SHVC streams with larger intra-periods also perform well under persistent bandwidth conditions due to reduced number of layer switching occurrences required.

The energy consumption at the client device also depends on the bandwidth fluctuation and SHVC encoding configuration. It can be observed that PW_t ratio is small under persistent network conditions and increases drastically under fluctuating

TABLE II
PERFORMANCE OF SHVC BIT STREAMING

Sequence	Intra-period=8			
	$f_{bw} = 5$		$f_{bw} = 20$	
	QL_t	PW_t	QL_t	PW_t
Race horses	0.296	1.503	0.341	1.581
People on street	0.337	1.651	0.300	1.336
Four people	0.413	1.519	0.462	1.349
Kimono	0.559	1.407	0.526	1.525
Bubbles	0.392	1.432	0.373	1.509
Average	**0.399**	**1.502**	**0.400**	**1.437**
Sequence	Intra-period=32			
	$f_{bw} = 5$		$f_{bw} = 20$	
	QL_t	PW_t	QL_t	PW_t
Race horses	0.283	1.584	0.327	1.429
People on street	0.316	1.558	0.281	1.571
Four people	0.410	1.608	0.459	1.611
Kimono	0.542	1.526	0.525	1.483
Bubbles	0.359	1.393	0.391	1.489
Average	**0.382**	**1.557**	**0.397**	**1.517**

bandwidths when the intra-period is large.

IV. CONCLUSION

The proposed adaptive streaming solution for SHVC standard shows that increasing random access points improves the video quality and power consumption performance under heavily fluctuating network conditions. Furthermore, less random access points are required when the network bandwidth is persistent. The future work will evaluate the proposed framework for multiple SHVC layers, with comparisons from other adaptive streaming technologies.

REFERENCES

[1] Cisco, "Cisco visual networking index: global mobile data traffic forecast update 2014-2019," Cisco, White Paper.

[2] G.J.Sullivan, J. Ohm, W-J Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[3] J. M. Boyce, Y. Ye, J. Chen, and A. K. Ramasubramonian, "Overview of SHVC: Scalable extensions of the high efficiency video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 20–34, July. 2015.

[4] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sept. 2009.

[5] W. Hamidouche, G. Cocherel, J. Le Feuvre, M. Raulet, and O. Dforges, "4K real-time video streaming with SHVC decoder and GPAC player," in *Proc. IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*. IEEE, 2014, pp. 1–2.

[6] T. Al Hadhrami, J. Nightingale, Q. Wang, C. Grecos, and N. Kehtarnavaz, "A simulator tool set for evaluating HEVC/SHVC streaming," *Proc.SPIE*, vol. 9400, pp. 9400 – 9400 – 10, 2015.

[7] J. Nightingale, Q. Wang, and C. Grecos, "Scalable HEVC (SHVC)-based video stream adaptation in wireless networks,," in *Proc. IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 2013, pp. 3573–3577.

[8] "openHEVC," https://github.com/OpenHEVC/openHEVC.

[9] "iPerf," https://iperf.fr/.

[10] "SHM4.1," https://hevc.hhi.fraunhofer.de/svn/svnSHVCSoftware/tags/SHM-4.1/.

[11] "tc," http://man7.org/linux/man-pages/man8/tc.8.html.

[12] "UPower, howpublished = https://upower.freedesktop.org/docs/Device.html," .