ADAPTIVE RATE CONTROL ALGORITHM FOR SHVC: APPLICATION TO HD/UHD

T. Biatek, W. Hamidouche†, J.-F. Travers‡ and O. Deforges†*

*IRT b<>com, Cesson-Sevigne, France
† IETR / INSA de Rennes, Rennes, France
‡ TDF, Cesson-Sevigne, France

ABSTRACT

Scalable video coding consists in compressing the video sequence into a layered bitstream where each layer refers to different spatial, temporal or quality representation of the video. Scalability enables compression gain compared to the simulcast encoding of layers thanks to inter-layer predictions. The scalable HEVC extension (SHVC) is the latest scalable technology promising up to 30% bitrate gains under the common test conditions, defined by JCT-VC. These conditions do not consider UHD and use fixed quantization step, which is not relevant in operational environment. In this paper, we propose an innovative adaptive rate control algorithm for SHVC. We consider HD as a base layer and UHD as an enhancement layer, with a constant global bitrate and a dynamic bitrate ratio adjustment between layers. The proposed algorithm is evaluated on a UHD data set where enables on average a BD-BR gain of 4.25% compared to a fixed-ratio encoding.

***Index Terms*—** HEVC, SHVC, UHD, HD, Rate-Control

1. INTRODUCTION

Scalable High Efficiency Video Coding [1] (SHVC) is the latest scalable video coding technology, released in October 2014 as the Annex H of the High Efficiency Video Coding (HEVC) [2] standard. With this extension, a video sequence can be compressed into spatial, temporal, SNR, bit-depth and color-gamut scalable layers. The Joint Collaborative Team on Video Coding (JCT-VC) announces up to 30% of bitrate reduction compared to an equivalent HEVC simulcast encoding [3], under the common test conditions (CTC) [4]. These conditions do not include Ultra High Definition (UHD) content and consider fixed quantization parameters (QP), which is irrelevant in operational environment. The previous scalable standard SVC [5], was published as an amendment of the AVC/H.264 standard [6]. The implementation changes required in AVC encoders for SVC migration and its late release led this technology to a limited industry adoption. However, the standard got a success in video-conferencing market where content are live-encoded and adaptively streamed over networks. For SHVC, JCT-VC focuses the standard development on simplicity, this way the standard provides only high level changes keeping the same encoding core. This quick release and the already existing real-time implementations [7] will enable faster industrial adoption. The Digital Video Broadcasting (DVB) group considers SHVC for insuring backward compatibility between UHD introduction phases. Adaptive Rate Control (ARC) consists in dynamically adjusting bitrate while encoding according to constraints which depend on the application. In Digital Terrestrial Television, channels are grouped into a multiplex which fits into an Ultra High Frequency (UHF) channel. While broadcasting, programmes are jointly encoded into a statistical pool in order to optimize bandwidth using statistical multiplexing (StatMux) techniques. In this case, the constraints may be the constant global bitrate related to UHF bandwidth or a minimal quality among programmes. For other applications such as Over-the-Top (OTT) content delivery services, bitrate constraints may come from the end user available bitrate which can vary in time. All these reasons motivate this work, and justify the relevance of proposing an ARC algorithm for SHVC. We apply SHVC to the spatial case with an High Definition (HD) Base Layer (BL) and a UHD Enhancement Layer (EL). The proposed algorithm dynamically adjusts the bitrate ratio between BL and EL to optimize the coding performance under constraint on the global bitrate .

This paper is organized as follows. In Section 2, the related work and the motivations are presented. Section 3 describes the proposed SHVC bitrate allocation algorithm. Section 4 describes experimental setup and provides an analysis and discussion of the results. Finally, Section 5 concludes this paper and draws up future work prospects.

2. RELATED WORK AND MOTIVATIONS

2.1. Rate-control in HEVC

During the HEVC standardization process, several RC algorithms were proposed. The Unified Rate Quantization (URQ) scheme [8] was firstly introduced in the HM7.1 (HEVC Test Model version 7.1) and replaced in the HM9.1 by a Rate-Lambda (R-λ) approach [9]. This model is based on a $R = \alpha \times \lambda^{\beta}$ function and composed of two mandatory bit allocation steps at Group Of Pictures (GOP) and Frame levels as well as one optional bit allocation step at the Coding Tree Unit (CTU) level. To achieve the targeted number of bits R, the related Lagrangian multiplier λ is computed and applied. Due to QP determination by λ value in HEVC Test Model, the appropriate QP is ascertained according to the selected λ. Other approaches were also proposed outside of JCT-VC. In [10], a ρ-Domain based approach is proposed for Low-Delay encoding configuration and then a Region-of-Interest-based (ROI) based bitrate allocation approach is proposed in [11]. In [12], an optimal ARC is proposed using an heavy pre-analysis over the video sequence. In [13], the bitrate is adaptively smoothed and allocated according to Virtual Buffer (VB) occupancy.

Table 1: SHVC scenarios and scalability impact

Scenario	SHVC	HEVC	Scalability impact
n°1	EL	EL Single layer equivalent	Gain
n°2	BL+EL	Simulcast equivalent	Gain
n°3	BL+EL	EL Single layer equivalent	Loss

2.2. Rate-control in SHVC

For SHVC, the URQ was firstly introduced in the SHM1.0 (SHVC Test Model version 1) and replaced by the R-λ approach in the SHM1.2 [14]. The HEVC R-λ is duplicated in each layer and several additional options are proposed such as separate R-λ models. In [15], an optimal bit allocation scheme for SHVC is provided. For each layer and its targeted bitrate, the budget allocated to each picture in the GOP is ascertained in an optimal way, through λ value and according to inter-layer dependencies. Several ARC-related contributions have been proposed for SVC. In [16], a streaming system using ARC is proposed where the amount of sent layers is adjusted depending on the estimated bandwidth. In [17], a StatMux using SVC fine-granular fidelity scalability is proposed. In this case the targeted bitrate of each programme is ascertained according to the estimated complexity and the bitrate is reached by refining the SVC bitstreams. All these contributions perform RC outside of the encoders, by truncating the scalable bitstreams to fit targeted bitrates. In [18], a complexity-aware ARC is proposed for ROI coding, the bitrate of layers are dynamically adjusted depending on the estimated computational complexity. In [19], a StatMux method using SVC for multicast application is proposed, programmes are jointly encoded separating BL and EL which are sent in different channels. Then the proposed joint rate controller adjusts bitrates according to estimated complexity. Eventually, an original contribution has been proposed in [20] based on game theory (GT) approach which optimizes performance by proposing an inner-layer frame-level bitrate allocation according to VB and bandwidth constraints. These contributions are efficient, but do not consider and exploit the impact of layers bitrate ratio over inter-layer (IL) tools and performances. Moreover, the already proposed method for SHVC [15] considers fixed targeted bitrate per layer, and is potentially compatible with our method since it improves the R-λ algorithm. Indeed, our proposed method consists in a novel ARC scheme for SHVC with the following features:

- accurate and constant BL+EL bitrate achievement,
- dynamic bit allocation within layers according to performance optimization on the specified ratio interval,
- light patch t ao the current R-λ algorithm with no significant additional complexity.

2.3. Impact of bitrate ratio over performance

In [21], an extended study on the impact of bitrate ratio over performance in SHVC is provided. The global bitrate and the bitrate ratio between BL and EL are defined as $R_G = R_{BL} + R_{EL}$ and $\tau = R_{BL}/R_G$. Where R_{BL} and R_{EL} are the bitrate in BL and EL, respectively. Four different global bitrate points are considered $R_G \in \{5, 10, 15, 20\} Mbps$ enabling to draw the Bjøntegaard Delta Bitrate (BD-BR) and BD-PSNR performance [22]. Then, 9 ratios $\tau \in \{0.1, 0.2, ..., 0.8, 0.9\}$ are applied to global bitrates R_G. The performance of SHVC can be evaluated through three scenarios, described in Table 1. In Scenario n°1, the UHD EL is compared to the EL-Equivalent Single Layer (SL) encoding and enables a gain thanks to IL predictions. In n°2 both BL and EL are compared to a Simulcast reference, a gain is expected since the EL is less expensive than EL in SL encoding. Eventually, both BL and EL are compared to EL-Equivalent SL HEVC encoding in n°3, in this case loss is expected because of the BL extra-cost. These scenarios have been evaluated with the latest reference software SHM9.0 [23] on the EBU UHD-1 data set which consists in ten UHD video sequences [24]. The targeted bitrates for both layers are reached with the R-λ algorithm in random access configuration. In Figure 1, we can observe the bitrate and PSNR losses of scenario n°3. Average bitrate and PSNR losses are recorded in ordinate and versus the 9 ratios in abscissa. We can notice that losses highly depend on the considered ratio and the video content. In scenario n°3, we can legitimately believe that losses can be reduced with a clever ratio adjustment. All these observations motivate this work since a significant potential gain is reachable.

3. PROPOSED METHOD

3.1. Parameters and indicators

In this paper, an ARC algorithm is proposed and applied to the HDTV/UHDTV 2x spatial scalability case, with a desired constant bitrate R_G. Let us define the framerate f and the total number of frames N_{Frames}. The initial targeted bits per picture in the sequence R_{AvgPic} is an encoding constant, defined as $R_{AvgPic} = R_G/f$. The RC algorithm requires three encoding indicators: the number of bits already used in the BL R_{BL}, the EL R_{EL} and the number of frames already coded N_{Coded}.

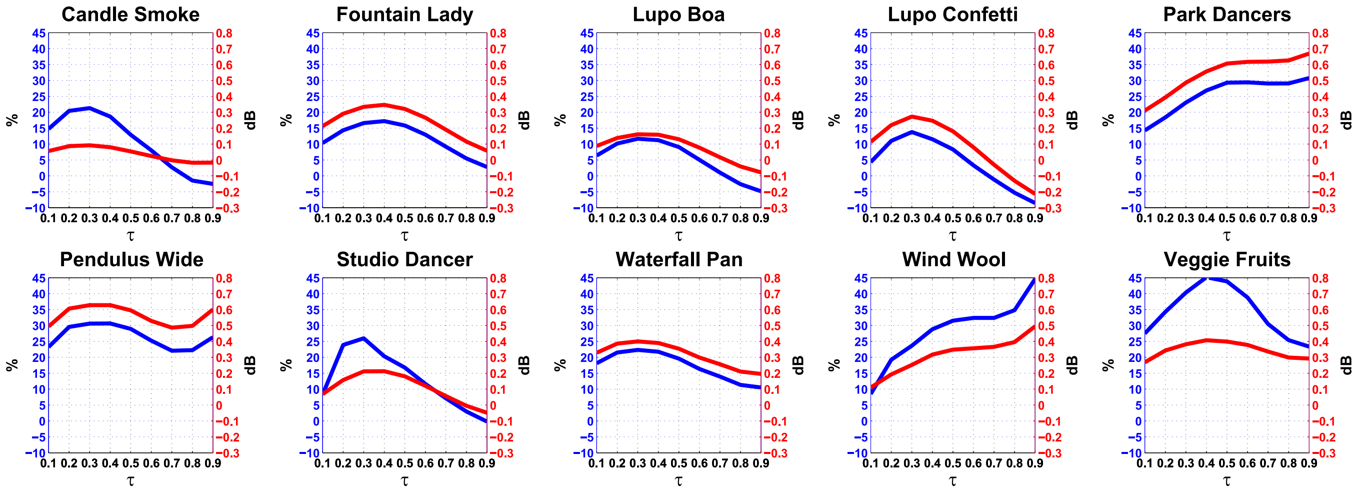

Fig. 1: Average bitrate and PSNR losses versus the bitrate ratio τ, for each sequence in the EBU-UHD1 data set

3.2. GOP-level adaptive rate control

The proposed algorithm substitutes the GOP-level bit allocation implemented in the SHM9.0 and keeps unchanged the underlying frame and CTU-level algorithms. Before each GOP encoding, the algorithm firstly updates the targeted bitrate in the upcoming GOP. The average number of bits per picture for the remaining frames T_{AvgPic} is firstly computed:

$$T_{AvgPic} = \frac{R_{PicAvg} \times (N_{Coded} + SW) - (R_{BL} + R_{EL})}{SW} \quad (1)$$

With SW the smoothing window width. Then, the global targeted number of bits for upcoming GOP is ascertained as:

$$T_{GOP} = T_{AvgPic} \times N_{GOP} \quad (2)$$

Where N_{GOP} is the number of frames in the upcoming GOP. In order to efficiently allocate this GOP budget within both layers, the optimal bitrate ratio τ_{opt} has to be determined. This is an optimization problem formulated as:

$$\tau_{opt} = \underset{\tau \in [\tau_{min}, \tau_{max}]}{\arg\max} \; G(\tau) \quad (3)$$

Where τ_{min} and τ_{max} are the minimal and maximal authorized ratios, respectively and G the estimated performance function. In order to smooth the bitrate variations, the optimal ratio value is clipped within the following interval:

$$\max(\tau_{last} \times 0.8, \tau_{min}) \leq \tau_{opt} \leq \min(\tau_{last} \times 1.2, \tau_{max}) \quad (4)$$

with τ_{last} the ratio used for the previous GOP. Once the optimal ratio ascertained, the related targeted number of bits for each layer T_{BL} and T_{EL} can be computed as:

$$T_{BL} = \tau_{opt} \times T_{GOP} \quad (5)$$

$$T_{EL} = (1 - \tau_{opt}) \times T_{GOP} \quad (6)$$

These bitrates feed the underlying frame and CTU level algorithms which accurately undertake targeted bitrates reaching. After each GOP, the parameters of performance function G are updated, as described in the following Section.

3.3. Estimation of gains function G

The performance function G has to be ascertained to solve optimization problem. In Figure 2, the BD-BR gains in the EBU data set are provided together with EL quality according to the ratio. We can notice that the EL quality in PSNR is highly correlated with the bitrate losses. In this work, we consider a fixed and centered ratio interval Φ defined as:

$$\Phi = [\tau_0 - 25\%, \tau_0 + 25\%] \quad (7)$$

We choose the centered value $\tau_0 = 1/2\sqrt{2}$ which is the commonly used value in the industry for HD to UHD, which is also close to the CTC average measured ratio (0.37) . Since Φ only concerns a short portion of the $[0, 1]$ interval, performing a parabolic estimation on Φ may be inaccurate because of GOP-level local fluctuation. This way, we consider a first order polynomial to locally represent the gain function on Φ:

$$G(\tau) = Q_{EL}(\tau) \triangleq \alpha \times \tau + \beta \quad (8)$$

After the encoding of the i-th GOP, the actual ratio τ_i and the EL global PSNR q_i are measured. The global PSNR is computed with usual weighting factors applied to average Y, U and V PSNR of the current GOP:

$$q_i = \frac{(6 \times Y_{PSNR} + U_{PSNR} + V_{PSNR})}{8} \quad (9)$$

Considering that a buffer of N previous samples (τ_i, q_i) is maintained, the α and β parameters can be estimated via the least squares method by solving the following equation system:

$$\begin{pmatrix} \sum \tau_i^2 & \sum \tau_i^1 \\ \sum \tau_i^1 & \sum \tau_i^0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \sum \tau_i q_i \\ \sum q_i \end{pmatrix} \quad (10)$$

Once the α and β parameters are ascertained, the optimization problem described in the Equation 3 can be easily solved and the appropriate ratio applied to the upcoming GOP.

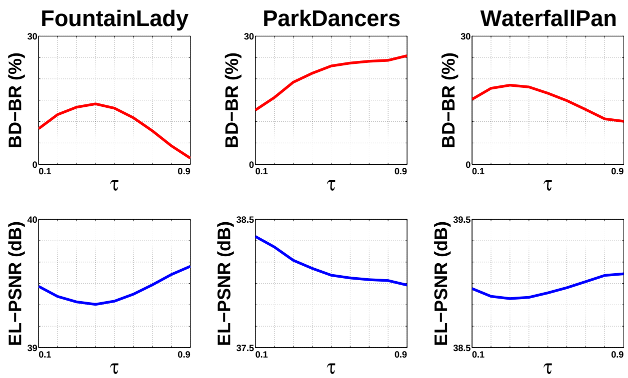

Fig. 2: BD-BR and EL-PSNR as a function of τ

4. RESULTS AND ANALYSIS

To evaluate the performance of the proposed method, the ARC algorithm has been implemented in the SHM9.0 and tested on the EBU UHD-1 data set which provides ten 3840x2160p sequences used as EL. The BL are built with the down-sampling tool provided in the reference software. The proposed algorithm is compared to the SL-Equivalent UHD HEVC encoding, and the resulting BD-BR performance is noted G_{ARC}. The SHM9.0 using constant $1/2\sqrt{2}$ ratio to build bitrate targets is also compared to SL-Equivalent UHD HEVC encoding, and the resulting BD-BR performance is noted G_{Ref}. The losses reduction of our method is also computed:

$$G_0 = \left(1 - \frac{G_{ARC}}{G_{Ref}}\right) \times 100 \qquad (11)$$

The BD-BR between proposed and reference algorithm is also computed and noted G_X. The whole results are recorded in Table 2 and the R-D curves of several sequences are provided in Figure 3. A positive value for G_{Ref}, G_{ARC} and G_X mean losses, while it means improvement for G_0.

Table 2: Performance of the proposed method

Sequence	G_{Ref}	G_{ARC}	G_0	G_X
Candle Smoke	18.97%	12.89%	32.03%	**-6.08%**
Fountain Lady	15.68%	16.16%	-3.09%	+0.48%
Lupo Boa	11.31%	10.46%	7.48%	**-0.85%**
Lupo Confetti	13.35%	11.60%	13.14%	**-1.75%**
Park Dancers	23.80%	22.22%	6.64%	**-1.58%**
Pendulus Wide	30.56%	26.49%	13.37%	**-4.07%**
Studio Dancer	18.92%	14.57%	22.98%	**-4.35%**
Waterfall Pan	20.72%	14.61%	29.50%	**-6.11%**
Wind Wool	19.64%	9.83%	49.94%	**-9.81%**
Veggie Fruits	31.33%	22.91%	26.90%	**-8.42%**
Average	20.43%	16.17%	19.89%	**-4.25%**

We can notice that the proposed ARC scheme improves performance compared to fixed-ratio encoding. The average losses are reduced by almost 20% in average, with a 15.5 standard deviation. All sequences are improved except for *Fountain Lady* which is slightly reduced by 3%. In this particular case, the fixed ratio value corresponds to the optimal ratio. The best optimization is reached in *Wind Wool* with 50% of losses reduction. The crossed BD-BR G_X is logically good with an average bitrate reduction of 4.25%, with a 9.8% peak and a -0.48% worst case. Regarding to computational complexity, the proposed method does not bring significant complexity since it substitutes the GOP-level existing method. The only additional complexity is the matrix inversion performed in Equation 10 by using a 2^{nd} order Gaussian elimination, this step is only performed once before each GOP encoding so around 93 times in the whole encoding (for this 10 seconds duration test set). The encoding time may change since a different ratio means different targeted bitrate for each layer. If a smaller ratio than reference $1/2\sqrt{2}$ is used, then a bigger bitrate is targeted in the UHD EL which leads to longer encoding time. If a bigger ratio than reference $1/2\sqrt{2}$ is selected, a lower bitrate would be targeted in the EL which results in a smaller encoding time.

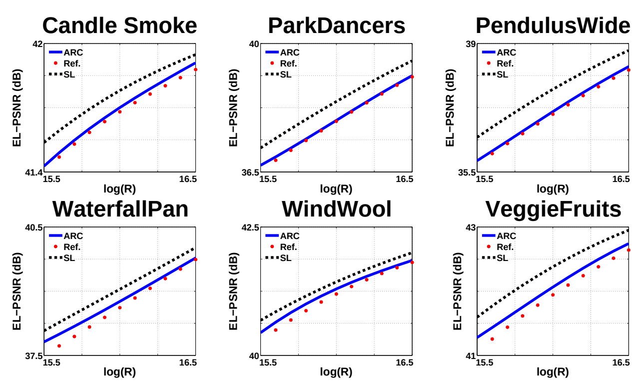

Fig. 3: R-D curves: ARC vs. Fixed-Ratio vs. SL

5. CONCLUSION

In this paper, an innovative ARC algorithm is proposed for SHVC. The experiments show an average reduction of 20% on compression losses for the EBU-UHD1 data set compared to the fixed-ratio reference coding, with the best and worst performance of 50% and -3%, respectively. In terms of BD-BR performance, this algorithm enables on average a significant gain of -4.25%. The proposed method substitutes the current GOP-level algorithm without bringing significant additional complexity. In future work, we will introduce to the encoder extra constraints such as upper and lower qualities and bitrates for both layers. These new parameters will help the algorithm to address new use-cases for SHVC such as OTT services, backward compatibility between DVB UHD-1 introduction phases or Hybrid Broadcast/Broadband services.

6. REFERENCES

[1] J. M. Boyce, Y. Ye, J. Chen, and A. K. Ramasubramonian, "Overview of SHVC: Scalable Extensions of the High Efficiency Video Coding (HEVC) Standard," *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, vol. 26, no. 1, pp. 20–34, Jan. 2016.

[2] "High Efficiency Video Coding," in *Rec. ITU-T H.265 and ISO/IEC 23008-2*, January 2013.

[3] V. Seregin and Y. He, "AHG Report: SHVC Software Development (AHG12)," in *Document JCTVC-R0012*. Sapporo, Japan, June 2014.

[4] ——, "Common SHM Test Conditions and Software Reference Configurations," in *Document JCTVC-Q1009*. Valencia, Spain, April 2014.

[5] H. Schwartz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[6] "Advanced Video Coding (MPEG-4 AVC)," in *Rec. ITU-T H.264 and ISO/IEC 14496-10*, May 2003.

[7] W. Hamidouche, M. Raulet, and O. Deforges, "4K Real-Time and Parallel Software Video Decoder for Multilayer HEVC Extensions," *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, vol. 26, no. 1, pp. 169–180, Jan. 2016.

[8] H. Choi, J. Nam, J. Yoo, D. Sim, and I. V. Bajic, "Rate control based on unified RQ model for HEVC," in *Document JCTVC-H0213*. San-Jose, USA, February 2014.

[9] B. Li, H. Li, L. Li, and J. Zhang, "λ-Domain Rate Control Algorithm for High Efficiency Video Coding," *IEEE Transactions on Image Processing*, vol. 23, no. 9, pp. 3841–3854, 2014.

[10] S. Wang, S. Ma, S. Wang, D. Zhao, and W. Gao, "Quadratic ρ-Domain Based Rate Control Algorithm for HEVC," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013.

[11] M. Meddeb, M. Cagnazzo, and B. Pesquet-Popescu, "Region-of-Interest-Based Rate Control Scheme for High-Efficiency Video Coding," *APSIPA Transactions on Signal and Information Processing*, vol. 3, no. 16, Dec. 2014.

[12] L. Sun, O. Au, C. Zhao, and F. Huang, "Rate Distorsion Modeling and Adaptive Rate Control Scheme for High Efficiency Video Coding (HEVC)," *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2014.

[13] J. Si, S. Ma, X. Zhang, and W. Gao, "Adaptive Rate Control for High Efficiency Video Coding," *Visual Communications and Image Processing (VCIP)*, Nov. 2012.

[14] B. Li, H. Li, and L. Li, "Rate control by R-lambda for SHVC," in *Document JCTVC-M0037*. Incheon, South Korea, April 2012.

[15] L. Li and H. Li, "λ Domain based Optimal Bit Allocation for Scalable High Efficiency Video Coding," *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2015.

[16] Y.-M. Hsiao, C.-H. Chen, J.-F. Lee, and Y.-S. Chu, "Designing and Implementing a Scalable Video-Streaming System Using an Adaptive Control Scheme," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 4, pp. 1314–1322, Nov. 2013.

[17] D. Jacobs, J. Barbarien, S. Tondeur, R. W. de Walle, T. Paridaens, and P. Schelkens, "Statistical Multiplexing Using SVC," *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2008.

[18] D. Grois and O. Hadar, "Complexity-Aware Adaptive Bit-Rate Control with Dynamic Pre-Processing for Scalable Video Coding," *IEEE International Conference on Multimedia and Expo (ICME)*, 2011.

[19] J. Jeong, S. Jeon, Y. Jung, and Y. Choe, "Statistical Multiplexing Using Scalable Video Coding for Layered Multicast," *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, May 2009.

[20] X. Wang, S. Kwong, L. Xu, and Y. Zhang, "Generalized Nash Bargaining Solution to Rate Control Optimization for Spatial Scalable Video Coding," *IEEE Transactions on Image Processing*, vol. 23, no. 9, pp. 4010–4021, 2014.

[21] T. Biatek, W. Hamidouche, J.-F. Travers, and O. Deforges, "Toward Optimal Bitrate Allocation in the Scalable HEVC Extension: Application to UHDTV," *IEEE International Conference on Consumer Electronics - Berlin (ICCE-Berlin)*, 2015.

[22] G. Bjøntegaard, "Calculation of Average PSNR Differences Between RD-Curves," in *Document VCEG-M33*, April 2001.

[23] J. Chen, J. Boyce, Y. Ye, and M. M. Hannuksela, "SHVC Test Model 9 (SHM 9) Introduction and Encoder Description," in *Document JCTVC-T1007*. Geneva, Switzerland, February 2015.

[24] EBU, "https://tech.ebu.ch/testsequences/uhd-1."