RATE CONTROL FOR LOSSLESS REGION OF INTEREST CODING IN HEVC INTRA-CODING WITH APPLICATIONS TO DIGITAL PATHOLOGY IMAGES

Victor Sanchez[1], Francesc Aulí-Llinàs[2], Rahul Vanam[3], and Joan Bartrina-Rapesta[2]

[1]Dept. of Computer Science, University of Warwick, Coventry, United Kingdom
[2]Dept. of Information and Communications Engineering, Universitat Autònoma de Barcelona, Spain
[3]Dept. of Electrical Engineering, University of Washington, Seattle, Washington, USA

ABSTRACT

This paper proposes a rate control algorithm for lossless region of interest (RoI) coding in HEVC intra-coding. The algorithm is developed for digital pathology images and allows for random access to the data. Based on an input RoI mask, the algorithm first encodes the RoI losslessly. According to the bit rate spent on the RoI, it then encodes the background by using rate control in order to meet an overall target bit rate. In order to increase rate control accuracy, the algorithm uses an *R*-*λ* model to approximate the slope of the rate-distortion curve, and updates any related model parameters during the encoding process. Random access is attained by coding the data using independent tiles. Experimental results show that the proposed algorithm attains the overall bit rate very accurately while providing lossless reconstruction of the RoI.

***Index Terms*—** HEVC, rate control, lossless RoI coding, medical images

1. INTRODUCTION

The introduction of high-throughput slide scanners has facilitated the widespread use of digital pathology images of microscope glass slides in clinical and research settings [1]. Recently, there has been an increasing interest in the area of telepathology in order to enable the examination of pathological specimens at a distance [2]. This would allow large national hospitals to provide telepathology services to regions where no pathology expertise exists, such as in some developing countries or isolated regions.

In telepathology systems with scarce transmission resources, it is advisable to focus transmission on those image regions used for analysis and diagnosis, also known as regions of interest (RoIs). In these situations, random access and RoI coding, in conjunction with lossless coding, provide an attractive solution for data access and transmission. For instance, pathologists with limited bandwidth connections using a remote image retrieval system may first access an RoI coded in a lossless manner, and then obtain a view of the entire image by accessing the background (BG), which may be coded in a lossy manner [3]. In these situations, rate control (RC) is an important tool that helps to deal with bit rate restrictions. Within this context, one of the most challenging tasks is achieving lossless RoI coding while maintaining an overall target bit rate by reducing the quality of the BG.

The High Efficiency Video Coding (HEVC) standard [4] has shown to provide an excellent performance for lossless coding of medical images, including lossless RoI coding in medical videos [5-7]. The HEVC reference software includes two different algorithms for RC. The first one is based on a quadratic model and the mean absolute difference (MAD) between the original and the reconstructed signals [8,9]. The second one is based on a rate-lambda (*R*-*λ*) model that takes into account the hierarchical coding structure [10]. Other proposals include textured and non-textured rate models [11], and an algorithm that processes RoIs and BG separately in a lossy manner by using two independent rate-distortion models [12].

Although current RC algorithms for HEVC provide a good performance, they are not designed for intra-coding of a single picture, and thus their accuracy in attaining a target bit rate is often low. In this paper, we then propose an RC algorithm for intra-coding of a single image that attains random access and lossless RoI coding by decreasing the BG quality according to a target bit rate. The proposed algorithm is based on the *R*-*λ* model of HEVC, which has proven to provide a better performance than the quadratic model [10]. It first allocates a bit budget to the BG according the cost of encoding the RoI and the overall bit rate. According to this allocated bit budget, it then computes the quantization parameter (QP) for each coding unit (CU) in the BG according to the *R*-*λ* model. In order to adapt to different images, the algorithm updates the parameters of the model with the encoding of CUs. In order to attain random access to the RoI and BG, the algorithm encodes CUs into independent tiles. The proposed RC algorithm is evaluated in a variety of pathology images at different bit rates. Evaluations show that the algorithm encodes the RoI losslessly while accurately attaining the overall target bit rate.

The rest of the paper is organized as follows. Section II briefly reviews the current HEVC RC algorithm based on the *R*-*λ* model. We describe our RC algorithm in Section III. Section IV presents the performance evaluation and Section V concludes this paper.

2. RATE CONTROL IN HEVC

Rate control in HEVC aims at allocating the appropriate number of bits and determining the quantization parameter of each picture and CU. The *R*-*λ* model in HEVC approximates the slope of the rate-distortion (RD) curve, denoted by λ, as follows:

$$\lambda = -\frac{\partial D}{\partial R} = \alpha \cdot R^{\beta} \tag{1}$$

where ∂ denotes a partial derivative, and α and β are parameters related to the RD characteristics of the sequence [10]. The algorithm comprises two main parts: 1) bit allocation, and 2) parameter adjustment to attain the target bit rate. It works at the group of pictures (GOP), picture and CU levels.

<u>GOP level</u>: based on a target bit rate R, the size of the GOP, the frame rate and a virtual buffer size, the algorithm computes the average number of bits per GOP.

<u>Picture level</u>: the target bit rate of each picture, $R_{Pic,}$ is computed according to the average allocated bits per picture and the

This work has been funded by the EU Marie Curie CIG Programme under Grant PIMCO and the Engineering and Physical Sciences Research Council (EPSRC), UK.

hierarchical level of the picture. Then, the R-λ model is used to compute the corresponding λ_{Pic} value. For inter-coded pictures, the model in (1) is used. Due to their high encoding cost, the λ_{Pic} value for intra-coded pictures is estimated using the mean absolute difference per pixel ($MADPP_{Pic}$) of the current picture as follows:

$$\lambda_{Pic} = \alpha \cdot (MADPP_{Pic} / R_{Pic})^{\beta} \tag{2}$$

where $MADPP_{Pic} = HAD_{Pic}/N_{Pic}$; HAD_{Pic} is the encoding cost computed as the sum of absolute differences between each CU in the current picture and their corresponding predictions in the horizontal and vertical directions (referred to as HAD cost [13]); and N_{Pic} is the number of pixels in the current picture.

<u>CU level</u>: for inter-pictures, the target bits of each CU are computed taking into account the bit budget assigned to the current picture, the number of already coded bits in the current picture and the weights of CUs. The latter are estimated as the MAD between the current CU and the same collocated CU in the previous coded picture at the same hierarchical level [10]. The model in (1) is then used to compute the corresponding λ value at the CU level, denoted by λ_{CU}, using the bit rate and $MADPP$ of each CU, denoted by R_{CU} and $MADPP_{CU}$, respectively [10].

For intra-pictures, the target bits of each CU are as follows:

$$T_{CU} = \frac{HAD_{CU} \cdot \omega}{HAD_{Pic} - HAD'_{Pic}} \tag{3}$$

where HAD_{CU} and HAD_{Pic} denotes the HAD cost of the current CU and current picture, respectively; HAD'_{Pic} denotes the HAD cost of the already coded CUs in the current picture; and ω denotes the number of bits left in the current picture budget weighted according to the number of CUs to be encoded and the number of bits already spent [10]. The corresponding λ_{CU} value is then computed using the model in (2) with values at a CU level, i.e., $MADPP_{CU}$ and $R_{CU} = T_{CU}/N_{CU}$, where N_{CU} is the number of pixels in the current CU.

In order to provide a good rate control performance with low computational complexity, at both the picture and CU levels, the QP values are determined as follows [14]:

$$QP = 4.2005 \ln \lambda + 13.7122 \tag{4}$$

A consistent quality is achieved by clipping all λ and QP values in a narrow range [10]. In the case of inter-coded pictures, the actual number of encoded bits and λ values are used to update the model parameters at the GOP, picture and CU levels [10].

3. PROPOSED RC ALGORITHM

The proposed RC algorithm is designed for a single intra-coded picture at a target bit rate R with lossless RoI coding and random access. The latter is achieved by coding the image using independent tiles. The algorithm aims at determining the quantization parameter of the whole BG and of each CU in the BG. It is based on the R-λ model of HEVC HM14.0+RExt 7.2 [15]. It uses an RoI mask as input to determine which CUs comprise the RoI and BG. Figure 1 illustrates the overall pipeline of the algorithm; note that it comprises a lossless mode and a lossy mode with rate control. For both modes, CUs are coded using tiling and intra-prediction based on sample-by-sample differential pulse code modulation (SbS-DPCM) [5,6]. All CUs in the RoI are first encoded losslessly by by-passing the quantization process and any other processing that affects the perfect reconstruction of the CUs. In this case, the residual signal is fed directly to the entropy coder.

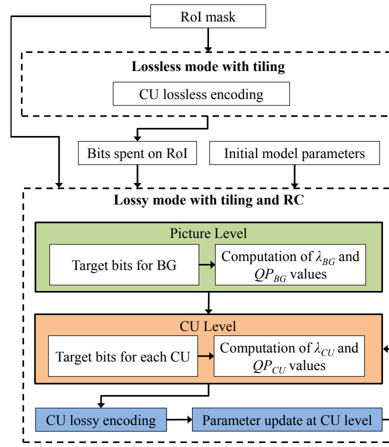


Fig. 1. Pipeline of the proposed RC algorithm.

All CUs in the BG are then encoded using the lossy mode with rate control. The RC algorithm works at the picture and CU levels.

At the *picture level*, the algorithm first estimates the target bit rate R_{BG} of the BG based on the overall target bit rate R and T_{RoI}, the number of bits spent to encode all the CUs in the RoI:

$$R_{BG} = \left(\left(R \cdot N_{Pic}\right) - T_{RoI}\right) / N_{BG} \tag{5}$$

where N_{Pic} is the number of pixels in the entire image and N_{BG} is the number of pixels in the BG. The λ and QP values for the BG, denoted by λ_{BG} and QP_{BG}, respectively, are then computed using the model in (1) and Eq. (4). The QP_{BG} value is later used at the CU level to clip the λ_{CU} values of each CU in a narrow range.

At the *CU level*, the target bits T_{CU} of each CU are first computed considering the HAD costs of only the CUs in the BG:

$$T_{CU} = \frac{HAD_{CU} \cdot \omega_{BG}}{HAD_{BG} - HAD'_{BG}} \quad \forall \text{ CU} \in \text{BG} \tag{6}$$

where HAD_{BG} denotes the HAD cost of the BG; HAD'_{BG} denotes the HAD cost of the already coded CUs in the BG; and ω_{BG} is the weighted number of bits left in the bit budget allocated to the BG:

$$\omega_{BG} = \widetilde{B}_{BG} + \left(\widetilde{B}_{BG} - \widetilde{T}_{BG}\right) \cdot \left(M / W_{BG}\right) \tag{7}$$

where $\widetilde{B}_{BG}$ denotes the actual number of bits left in the BG budget computed using the actual number of bits spent on the already coded BG CUs; $\widetilde{T}_{BG}$ denotes the number of bits left in the BG budget computed using the target bits (T_{CU}) of the already coded BG CUs; M denotes the number of BG CUs yet to be encoded; and W_{BG} is a constant used to guarantee that the BG budget is respected. The corresponding λ_{CU} is then computed by (1) using the bit rate of each CU, R_{CU}. The QP value for each CU, denoted by QP_{CU}, is computed using Eq. (4).

3.1 Model parameter updating

Unlike the current RC algorithm, which does not update the model parameters for intra-coded pictures, the proposed algorithm updates these parameters. However, for a single intra-coded picture, updating the parameters at the GOP and picture level is not feasible, which limits the adaptability of the model. In order to accurately attain the overall bit rate, it is important to update parameters at the CU level during the encoding process of the BG. Therefore, the model parameters are updated at the CU level using the actual bit rate R_{CU} and the actual λ_{CU} value of each CU. Specifically, the α and β values are updated as the geometric mean of the α and β values of the already coded CUs in the BG. After the encoding of each CU, these parameters are updated as follows [10]:

$$\lambda_{CUr} = \alpha_{old} \cdot (R_{CUr})^{\beta_{old}} \quad (8)$$

$$\alpha_{new} = \alpha_{old} + \delta_{\alpha} \cdot (\ln \lambda_{CU} - \ln \lambda_{CUr}) \cdot \alpha_{old} \quad (9)$$

$$\beta_{new} = \beta_{old} + \delta_{\beta} \cdot (\ln \lambda_{CU} - \ln \lambda_{CUr}) \cdot \ln R_{CUr} \quad (10)$$

where λ_{CU} is the estimated value for the current CU computed by the model in (1); R_{Cur} is the actual bit rate used to encode the current CU; λ_{CUr} is the value calculated using the actual number of encoded bits of the current CU; and δ_{α} = 0.1 and δ_{β} = 0.05 are constants that control the updating process [10].

3.2 Background quality consistency

Finally, in order to attain quality consistency in the whole BG, it is important to clip the computed λ and QP values in a narrow range. At the *picture level*, the QP_{BG} value is guaranteed that:

$$QP_{\min} \le QP_{BG} \le QP_{\max} \quad (11)$$

where $QP_{\min}$ and $QP_{\max}$ are the minimum and maximum QP values allowed by the encoder, respectively.

The current RC algorithm clips the λ_{CU} and QP_{CU} values in a narrow range determined by the QP value of the previously coded CU, which is usually spatially adjacent to the current CU. With lossless RoI coding, some of the CUs in the BG may not be spatially adjacent, even if the proposed algorithm encodes them sequentially after coding all the RoI CUs. Therefore, clipping the current λ_{CU} and QP_{CU} values based only on the QP value of the previously coded BG CU may result in CUs encoded at very low or high qualities, particularly those CUs surrounding the RoI. In order to help avoiding this, at the *CU level*, the proposed RC algorithm clips λ_{CU} and QP_{CU} values in a narrow range determined by the average QP value of the BG:

$$\lambda_{\min} \le \lambda_{CU} \le \lambda_{\max} \quad (12)$$

$$\overline{QP}_{BG} - \sigma \le QP_{CU} \le \overline{QP}_{BG} + \sigma \quad (13)$$

with $\lambda_{\min}$ and $\lambda_{\max}$ calculated as follows:

$$\lambda_{\min/\max} = e^{((\overline{QP}_{BG} -/+ \sigma) - 13.7122)/4.2005} \quad (14)$$

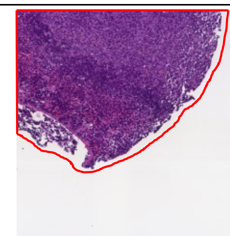
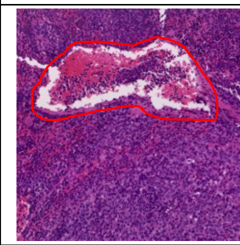
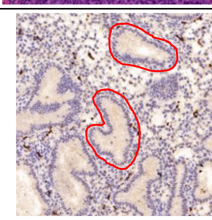
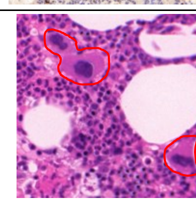
where σ = 4 is a constant that allows to accommodate for quality differences between the overall quality of the BG and that of those CUs with high or low encoding costs, i.e., high or low $MADPP_{CU}$ values; and $\overline{QP}_{BG}$ is the arithmetic mean of the QP_{CU} vales of the m previously coded CUs in the BG:

$$\overline{QP}_{BG} = \begin{cases} QP_{BG} & \text{if CU} = 1 \\ \sum_{codedBGCUs} QP_{CU} / m & \text{if CU} > 1 \end{cases} \quad (15)$$

4. EXPERIMENTAL EVALUATIONS

The proposed rate control algorithm is implemented in the reference software HM14.0 with range extensions (RExt 7.2), and evaluated for the three different RoI coding cases, which are likely to be encountered in telepathology. The first case is lossless coding of the entire depicted specimen with lossy coding of the empty BG. The second case is lossless coding of a single RoI with lossy coding of the rest of the image. The last case is lossless coding of more than one RoI with lossy coding of the rest of the image. These three cases are evaluated with five different pathology images from the Center for Biomedical Informatics and Information Technology of the US National Cancer Institute and the University Hospital Coventry and Warwickshire, UK.

Table 1.Information about the test images and corresponding RoIs.

Image (resolution:bpp)	CU/tile size	No. of RoIs	RoI size	RoI location
1.PAT1 (1920×1920:24)	64×64	1	54.67%	
2.PAT2 (1280×1280:24)	64×64	1	18.25%	
3.PAT3 (1024×1024:24)	32×32	2	12.40%	
4.PAT4 (448×448:24)	16×16	2	11.10%	
5.PAT5 (448×448:24)	16×16	1	9.44%	

All test images are intra-coded as a single RGB picture using SbS-DPCM in 4:4:4 format. SbS-DPCM is implemented in all prediction block sizes with a minimum size of 4×4 and a maximum size equal to the CU size. In order to attain random access, images are encoded using independent tiles, whose sizes are equal to the CU sizes. The resulting bit-stream can then be easily split into two sub-streams, one for RoI and BG, by altering some flags in the headers [16,17]. Random access is then attained with a granularity level equal to the CU size. Table 1 tabulates the information about the test images. All RoI masks are computed *a priori* by manually delineating the RoIs. In this work, any CU that contains RoI pixels

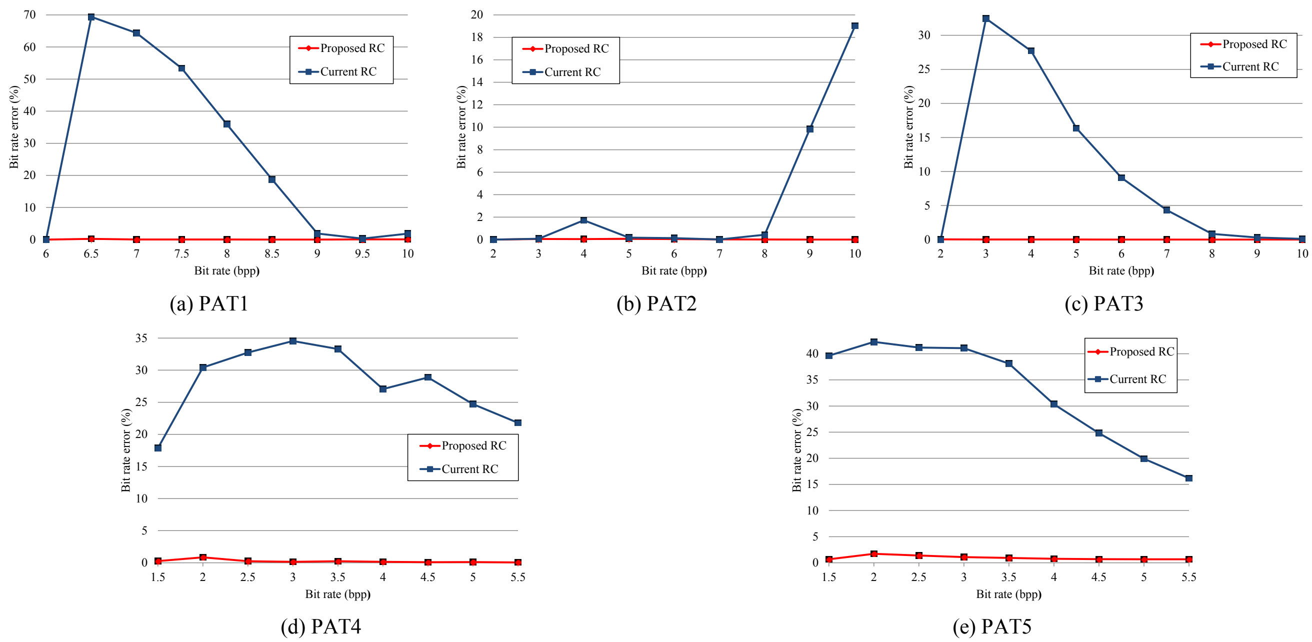


Fig. 2. Bit rate error, in percentage, of the proposed and current RC algorithms for different bit rates and test images. In all cases, the RoI is encoded losslessly. Bit rates shown correspond to overall bit rates (RoI + BG).

is considered as part of the RoI.

The performance of the proposed RC algorithm is compared to the current RC algorithm as implemented in HM14.0+RExt 7.2. In order to provide a fair comparison, the current RC algorithm is modified to allow for lossless RoI coding. This is done by coding all RoI CUs losslessly first, followed by BG CUs using the current RC algorithm. Figure 2 plots the accuracy of both algorithms, in terms of the of bit rate error (BRE), in percentage, at different bit rates. These results show that the proposed RC algorithm attains the overall bit rate very accurately with BREs close to zero. This is expected, as the algorithm updates the model parameters during the encoding process. As a consequence, the initial values of α and β are not critical, as this updating process helps adapting the model to the characteristics of the image and results in an accurate approximation of the RD curve. It is important to mention that in images PAT1 and PAT2, the RoI requires just over 6bpp and 2bpp, respectively, to be encoded losslessly. Therefore, the lowest bit rate shown in Fig. 2(a)-(b) is that needed to encode the RoI losslessly with no BG. The target bits for the BG are thus equal to zero for these two bit rates and both RC algorithms encode all CUs in the BG as zero-valued data. As a consequence, the BRE of both algorithms is zero for these lowest bit rates.

It is interesting to note that, in general, the accuracy of the current RC algorithm tends to increase as the bit rate increases, despite the fact that the model parameters are kept constant. This means that the model in (2) represents the RD curve with relatively high accuracy for high bit rates. The performance is, however, very poor for low bit rates. It is also worth noting that in the proposed RC algorithm, the bit rate errors tend to be smaller for images with large BGs, e.g., PAT1, as opposed to images with small BGs, e.g., PAT5. This may be explained by the fact that large BGs usually comprise a large number of CUs, which allows the algorithm to better adjust the R-λ model to the characteristics of the image.

Table 2 tabulates PSNR values for the BG of images PAT1, PAT2 and PAT3 at the three bit rates where the BRE of both algorithms is the most similar. Note that for these bit rates, both RC algorithms go over the bit rate; however, the BREs of the current RC algorithm are, in general, higher and thus more bits are spent in the BG. Consequently, the quality attained by this algorithm is expected to be higher than that attained by the proposed one. In cases where both algorithms attain a similar BRE, e.g., PAT2 at 3.0 bpp, the average quality difference between them is only 0.07 dB. This shows that both algorithms can encode the BG with similar qualities; however the rate control mechanism of the proposed one is far more accurate, as shown in Fig. 2.

Table 2. PSNR values (dB) of the BG region at those bitrates where the proposed and current RC algorithms attain the most similar accuracy.

Image	Rate-bpp	Proposed RC				Current RC			
		BRE	R	G	B	BRE	R	G	B
PAT1	9.0	0.013%	55.13	54.40	54.38	1.902%	56.51	55.24	55.24
	9.5	0.058%	58.29	57.06	57.04	0.294%	60.35	58.03	58.01
	10.0	0.073%	65.57	65.57	62.56	1.860%	66.70	61.85	61.84
PAT2	3.0	0.055%	33.13	33.37	33.29	0.077%	33.20	33.44	33.36
	6.0	0.044%	44.23	44.05	43.97	0.130%	44.75	44.50	44.42
	7.0	0.021%	47.64	47.31	47.23	0.010%	47.08	46.78	46.69
PAT3	8.0	0.002%	47.17	46.48	46.40	0.841%	48.37	47.63	47.61
	9.0	0.002%	49.52	48.68	48.64	0.305%	50.45	49.55	49.55
	10.0	0.004%	51.70	50.65	50.65	0.106%	52.61	51.45	51.50

5. CONCLUSIONS

This paper presents a rate control algorithm for lossless RoI intra-coding of single images with random access in HEVC. The algorithm, which is based on an R-λ model, processes coding units in the RoI losslessly followed by those in the background by using rate control. It attains an overall bit rate very accurately thanks to an updating process that adjusts the RD model parameters during the encoding process. Evaluations over various pathology images show the high accuracy of the rate control mechanism. This algorithm is also useful for other medical and natural images, as well as screen content material.

6. REFERENCES

[1] A. Madabhushi, "Digital pathology image analysis: Opportunities and challenges (editorial)," *Imag. Med.*, vol. 1, no. 1, pp. 7-10, 2009.

[2] M. May, "A better lense on disease," *Scientific American*, vol. 302, pp. 74-77, 2012.

[3] V. Sanchez, "Joint Source/Channel Coding for Prioritized Wireless Transmission of Multiple 3D Regions of Interest in 3D Medical Imaging Data," *IEEE Transactions on Biomedical Engineering*, vol. 60. no. 2, pp. 397-405, 2013.

[4] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard*," IEEE Trans. on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 2012 1649-1668 December 2012.

[5] V. Sanchez, and J. Bartrina-Rapesta, "Lossless compression of medical images based on HEVC intra coding," in *Proc. IEEE Int. Conf. Acoustics, Speech Signal Processing (ICASSP)*, pp. 6622-6626, May 2014.

[6] V. Sanchez, F. Aulí-Llinàs, J. Bartrina-Rapesta, and J. Serra-Sagristà, "HEVC-based Lossless Compression of Whole Slide Pathology Images," in *Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, December 2014, *in press*.

[7] H. Chen, G. Braeckman, S.M. Satti, P. Schelkens, and A. Munteanu, "HEVC-based video coding with lossless region of interest for telemedicine applications," in *Proc. 20th International Conference on Systems, Signals and Image Processing (IWSSIP)*, pp. 129-132, July 2013.

[8] H. Choi, J. Nam, J. Yoo, and D. Sim, "Rate control based on unified RQ model for HEVC," in *JCT-VC H0213, 8th meeting of Joint Collaborative Team on Video Coding of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11*, San Jose, CA, USA, 2012.

[9] ——, "Improvement of the rate control based on pixel-based URQ model for HEVC," in *JCT-VC I0094, 9th meeting of Joint Collaborative Team on Video Coding of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11*, Geneva, CH, 2012.

[10] B. Li, H. Li, L. Li, and J. Zhang, "Rate control by R-lambda model for HEVC," in *JCT-VC K0103, 11th meeting of Joint Collaborative Team on Video Coding of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11*, Shanghai, CN, 2012.

[11] B. Lee, M. Kim, and T. Nguyen, "A frame-level rate control scheme based on texture and non-texture rate models for High Efficiency Video Coding," *IEEE Trans. Circuits Syst. Video Technol.*, pp. 1-14, 2013.

[12] M. Meddeb, M. Cagnazzo, and B. Pesquet-Popescu, "Region-of-interest based rate control scheme for high efficiency video coding," in *Proc. IEEE Int. Conf. Acoustics, Speech Signal Processing (ICASSP)*, pp. 7338-7342, May 2014.

[13] Y. Kim, D. Jun, S. Jung, and J. Choi, "A fast intra prediction method using Hadamard transform in high efficiency video coding," in *Proc. Visual Information Processing and Communication III*, pp. 270-280, February 2012.

[14] B. Li, D. Zhang, H. Li, and J. Xu, "QP determination by lambda value," in *JCT-VC I0426, 9th meeting of Joint Collaborative Team on Video Coding of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11*, Geneva, CH, 2012

[15] HM14.0+RExt 7.2 software. [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-14.0+RExt-7.2/

[16] P. Amon, M. Sapre, and A. Hutter, "Compressed domain stitching of HEVC streams for video conferencing applications," in *Proc. 2012 19th International Packet Video Workshop (PV)*, pp.36-40, May 2012.

[17] C. Feldmann, C. Bulla, and B. Cellarius, "Efficient Stream-Reassembling for Video Conferencing Applications using Tiles in HEVC," in *Proc. 2013 Fifth International Conferences on Advances in Multimedia (MMEDIA)*, pp. 130-135, April 2013.