**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC1/SC29/WG11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 MPEG2012/M26638**

**October 2012, Shanghai, China**

|  |  |
| --- | --- |
| **Source** | **Kwangwoon University (KWU) and Simon Fraser University (SFU)** |
| **Status** | **Proposal** |
| **Title** | **3D-HEVC – Improvement of the rate control for 3D multi-view video coding** |
| **Author** | **Woong Lim (limwoong@kw.ac.kr)**  **Hyomin Choi (hyomin06@kw.ac.kr)**  **Junghak Nam (qejixfyza@kw.ac.kr)**  **Hyunho Jo (idjhh@kw.ac.kr)**  **Donggyu Sim (dgsim@kw.ac.kr)**  **Ivan V. Bajić (ibajic@ensc.sfu.ca)** |

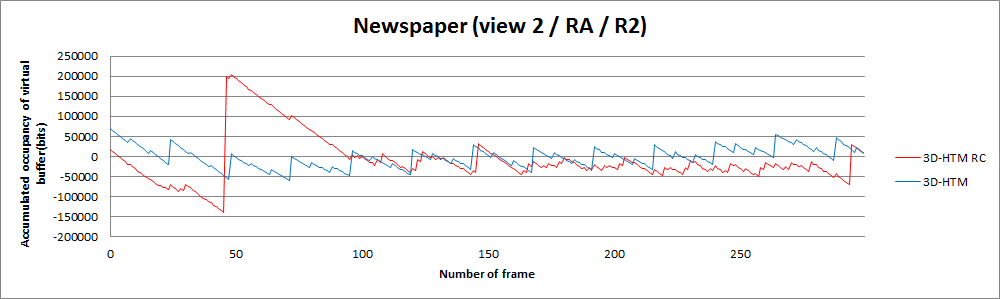
# Abstract

In the last meeting in Stokholm, we proposed a rate control algorithm for multi-view video coding [1]. That contribution presented a rate control scheme for multi-view video coding based on the unified rate-quantization (URQ) model for HEVC described in JCTVC-H0213 [2] and JCTVC-I0094 [3]. In the contribution, there were two primary algorithms for extended view: (1) initial frame level QP setting of the first frame of a GOP of an extended view, and (2) inter-view MAD prediction by using the depth map. However, a problem was observed for some sequences where sudden spikes in the number of generated bits occured in extended views. This problem did not affect much the average performace of rate control, but still caused a noticeable fluctuation in the number of generated bits on a frame by frame basis in extended views. To mitigate this problem, we propose a new method for initial QP setting for the first frame of a GOP of extended views. The proposed rate control scheme, incorporating the new initial QP setting, is implemeted on 3D-HTM4.0.1. To evaluate the proposed scheme, 3D-HTM4.0.1, which uses hierarchical QP setting, is used for the constant bit rate (CBR) case. In the previous contribution, we found that the accuracy of our proposed multi-view rate control was over 99.11%, which came with a 2.1dB of average PSNR degradation. By using the new improved rate control scheme, the average bitrate accuracy is 99.20% with only about 1.22dB average PSNR degradation.

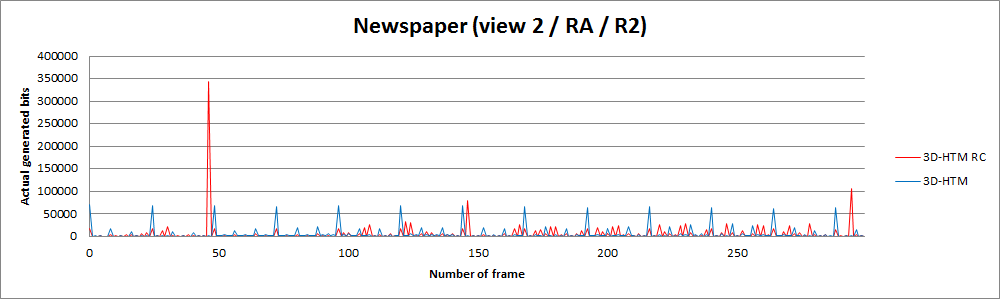
# Introduction

A rate control algorithm for HEVC, described in JCTVC-H0213, was adopted in the reference HM software at the 8th San Jose meeting in February 2012. Some improvements were reported in JCTVC-I0094. While the approach from JCTVC-H0213 and JCTVC-I0094 can be applied to HEVC-based multi-view encoding, some modifications are necessary in order to adapt the rate control procedure to the multi-view scenario. In the last meeting in Stokholm, we proposed rate control scheme for multi-view video coding [1] with consideration of the following two main issues.

First, we considered how to set the initial QP value of the first frame of the extended views. Because the extended view has at least one reference frame at any point in time, QP initialization can be conducted using the QP of the inter-view reference frame at the frame level. Furthermore, for stable generation of bits, we proposed inter-view MAD prediction for extended views. In single-view coding, the temporal MAD factor is used to predict the complexity of texture for the current frame. In single-view coding, to predict the complexity of texture, there is no other choice but to use the previously encoded frame along the temporal axis to calculate temporal MAD. But in multi-view coding, the previously encoded frame(s) in neighbouring view(s) can also be employed. In addition, depth map can be used to find the corresponding position of the current unit block in the neighbouring view. While the proposed method was able to control the aggregate bitrate fairly well, there was an occasional problem in some sequences that caused sudden spikes in the number of generated bits at some frames in extended views, as shown in Fig. 1.



(a)



(b)

Figure 1. Frame-wise (a) accumulated occupancy of the virtual buffer and (b) actual generated bits for ‘Newspaper’ sequence (view 2, R2)

To remove this problem, we corrected frame-level QP initialization of the first frame of the GOP of an extended view to reference the QP from the base view and use the pre-defined view layer QP offset.

# Initial QP decision for the first frame of a GOP in an extended view

In the previous rate control method [1], the QP for the first frame of a GOP was initialized by the QP of the first frame of the previous GOP, for both base view and extended view. While the strategy is reasonable for the base view, in the case of an extended view, a better strategy would be to rely on the QP of the corresponding frame in the base view.

At the beginning of encoding, the first frame of an extended view is an inter-view frame. This means that the first frame of an extended view is not coded as an intra-frame but as an inter-frame. At the frame level, the QPs between the base view and an extended view are redundant and to maintain a similar quality between the base view and an extended view, *ViewLayerQPOffset* is pre-defined in the configuration file. We can simply use this *ViewLayerQPOffset* for initialization of the frame-level QP of the first frame of an extended view in the GOP.

 (1)

*ViewLayerQPOffset* indicates the QP offset between view layers.  and  are QPs of the first frame of the extended view and the base view, respectively. By referencing the QP of the first frame of the GOP from the base view as above, the bit rate and the quality of the extended view can be made more stable.

# Experimental results

The proposed rate control for multi-view video coding is tested under the CBR condition. Table 1 presents the delta bitrate to show the overall performance. The proposed rate control is evaluated under the same conditions of the common random access condition of the current 3D multi-view coding. To show the difference in the number of generated and target bits, we use % kbps error defined by formula (2).

 (2)

Table 1. Performance of view-wise actual generated bits of 3D-HTM rate control and the existing 3D-HTM

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **sequence** | **target total** | **Target rate V0** | **Target rate V1** | **Target rate V2** | **view0** | **view1** | **view2** | **% kbps error view0** | **% kbps error view1** | **% kbps error view2** | **overall kbps error(%)** |
| **Balloons** | 1527.00 | 876.00 | 311.00 | 340.00 | 872.84 | 310.65 | 340.75 | 0.36 | 0.11 | 0.22 | 0.18 |
| 797.00 | 482.00 | 151.00 | 164.00 | 479.50 | 148.60 | 163.41 | 0.52 | 1.59 | 0.36 | 0.69 |
| 448.00 | 278.00 | 81.00 | 89.00 | 276.02 | 79.31 | 87.67 | 0.71 | 2.09 | 1.49 | 1.12 |
| 265.00 | 168.00 | 46.00 | 51.00 | 166.92 | 45.75 | 50.13 | 0.64 | 0.55 | 1.71 | 0.83 |
| **Kendo** | 1495.00 | 844.00 | 314.00 | 337.00 | 832.41 | 307.95 | 331.78 | 1.37 | 1.93 | 1.55 | 1.53 |
| 759.00 | 451.00 | 149.00 | 159.00 | 446.39 | 147.66 | 157.78 | 1.02 | 0.90 | 0.77 | 0.94 |
| 423.00 | 257.00 | 81.00 | 85.00 | 256.20 | 80.56 | 84.66 | 0.31 | 0.54 | 0.41 | 0.37 |
| 250.00 | 155.00 | 46.00 | 49.00 | 154.86 | 45.96 | 48.99 | 0.09 | 0.08 | 0.02 | 0.07 |
| **Newspaper** | 1685.00 | 934.00 | 383.00 | 368.00 | 924.76 | 379.81 | 360.92 | 0.99 | 0.83 | 1.92 | 1.16 |
| 839.00 | 491.00 | 180.00 | 168.00 | 486.02 | 175.68 | 164.10 | 1.02 | 2.40 | 2.32 | 1.57 |
| 453.00 | 270.00 | 95.00 | 88.00 | 266.27 | 92.56 | 86.02 | 1.38 | 2.57 | 2.25 | 1.80 |
| 260.00 | 157.00 | 54.00 | 49.00 | 152.44 | 52.37 | 47.77 | 2.90 | 3.01 | 2.51 | 2.85 |
| **Gtfly** | 4949.00 | 3670.00 | 629.00 | 650.00 | 3650.33 | 622.29 | 649.56 | 0.54 | 1.07 | 0.07 | 0.54 |
| 2091.00 | 1611.00 | 235.00 | 245.00 | 1602.14 | 234.06 | 242.14 | 0.55 | 0.40 | 1.17 | 0.61 |
| 976.00 | 755.00 | 107.00 | 114.00 | 751.92 | 107.22 | 113.44 | 0.41 | 0.20 | 0.49 | 0.35 |
| 474.00 | 366.00 | 52.00 | 56.00 | 364.24 | 52.23 | 56.47 | 0.48 | 0.44 | 0.84 | 0.22 |
| **PoznanHall2** | 1366.00 | 809.00 | 271.00 | 286.00 | 802.45 | 257.47 | 291.30 | 0.81 | 4.99 | 1.85 | 1.08 |
| 570.00 | 344.00 | 111.00 | 115.00 | 341.25 | 111.43 | 115.51 | 0.80 | 0.39 | 0.44 | 0.32 |
| 300.00 | 178.00 | 60.00 | 62.00 | 181.27 | 62.00 | 63.34 | 1.83 | 3.33 | 2.16 | 2.20 |
| 167.00 | 100.00 | 33.00 | 34.00 | 99.54 | 33.09 | 34.27 | 0.46 | 0.28 | 0.80 | 0.06 |
| **PoznanStreet** | 3913.00 | 2339.00 | 707.00 | 867.00 | 2332.93 | 695.76 | 854.78 | 0.26 | 1.59 | 1.41 | 0.75 |
| 1355.00 | 890.00 | 209.00 | 256.00 | 885.74 | 196.04 | 249.00 | 0.48 | 6.20 | 2.73 | 1.79 |
| 605.00 | 415.00 | 86.00 | 104.00 | 413.51 | 86.70 | 104.72 | 0.36 | 0.81 | 0.70 | 0.01 |
| 302.00 | 212.00 | 42.00 | 48.00 | 211.53 | 41.53 | 47.74 | 0.22 | 1.12 | 0.54 | 0.40 |
| **Dancer** | 6539.00 | 4593.00 | 1002.00 | 944.00 | 4578.18 | 988.16 | 925.57 | 0.32 | 1.38 | 1.95 | 0.72 |
| 2736.00 | 1996.00 | 381.00 | 359.00 | 1990.06 | 381.46 | 360.73 | 0.30 | 0.12 | 0.48 | 0.14 |
| 1256.00 | 917.00 | 174.00 | 165.00 | 914.53 | 175.06 | 165.68 | 0.27 | 0.61 | 0.41 | 0.06 |
| 592.00 | 429.00 | 83.00 | 80.00 | 427.83 | 83.49 | 80.64 | 0.27 | 0.59 | 0.80 | 0.01 |
|  |  |  |  |  |  |  | **Avg** | 0.70 | 1.43 | 1.16 | 0.80 |

Table 1 shows the percentage error of the proposed rate control given the target bitrates produced by the existing 3D-HTM. The average percentage error across all views and sequences is about 0.80%, meaning that the average accuracy is 99.20%.

However, rate control comes with some cost in terms of PSNR degradation. To show the PSNR degradation from the proposed rate control, we computed the ΔPSNR as

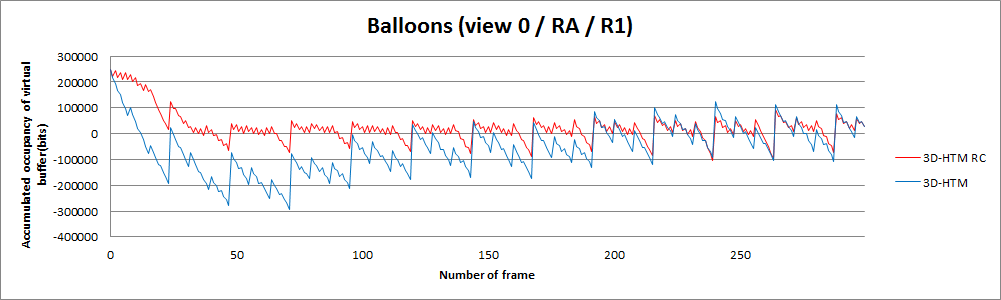
 (3)

Table 2. Performance of view-wise quality of 3D-HTM rate control and the existing 3D-HTM

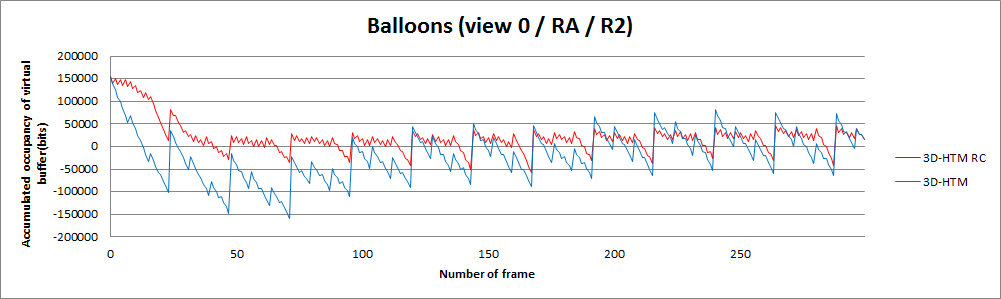
|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **sequence** | **PSNR V0** | **PSNR V1** | **PSNR V2** | **view0** | **view1** | **view2** | **Δ PSNR V0** | **Δ PSNR V1** | **Δ PSNR V2** | **average Δ PSNR** |
| **Balloons** | 42.51 | 41.36 | 41.12 | 43.10 | 42.29 | 41.96 | 0.60 | 0.93 | 0.84 | 0.79 |
| 39.80 | 38.69 | 38.52 | 41.07 | 39.99 | 39.74 | 1.27 | 1.30 | 1.22 | 1.26 |
| 36.37 | 35.39 | 35.30 | 38.51 | 37.40 | 37.19 | 2.13 | 2.01 | 1.89 | 2.01 |
| 33.87 | 32.80 | 32.58 | 35.68 | 34.67 | 34.47 | 1.81 | 1.87 | 1.89 | 1.85 |
| **Kendo** | 41.97 | 40.64 | 40.59 | 43.62 | 42.75 | 40.17 | 1.65 | 2.11 | -0.42 | 1.11 |
| 39.52 | 38.38 | 38.33 | 41.46 | 40.43 | 40.17 | 1.95 | 2.05 | 1.84 | 1.94 |
| 37.09 | 36.22 | 36.09 | 38.99 | 38.01 | 37.81 | 1.90 | 1.79 | 1.72 | 1.80 |
| 34.68 | 33.92 | 33.78 | 36.38 | 35.47 | 35.32 | 1.70 | 1.55 | 1.54 | 1.60 |
| **Newspaper** | 40.17 | 39.18 | 38.80 | 41.38 | 40.25 | 39.87 | 1.21 | 1.07 | 1.07 | 1.12 |
| 37.35 | 36.16 | 35.92 | 38.98 | 37.85 | 37.48 | 1.63 | 1.68 | 1.57 | 1.63 |
| 34.87 | 33.79 | 33.58 | 36.46 | 35.34 | 35.04 | 1.59 | 1.55 | 1.46 | 1.54 |
| 31.96 | 30.86 | 30.56 | 33.93 | 32.86 | 32.55 | 1.97 | 2.00 | 1.99 | 1.98 |
| **Gtfly** | 39.44 | 38.74 | 38.76 | 40.19 | 39.68 | 39.64 | 0.75 | 0.94 | 0.88 | 0.86 |
| 36.61 | 36.26 | 36.15 | 37.78 | 37.43 | 37.38 | 1.17 | 1.16 | 1.23 | 1.19 |
| 34.60 | 34.46 | 34.42 | 35.60 | 35.36 | 35.31 | 1.00 | 0.90 | 0.89 | 0.93 |
| 33.03 | 32.89 | 32.78 | 33.46 | 33.29 | 33.23 | 0.44 | 0.40 | 0.45 | 0.43 |
| **PoznanHall2** | 41.54 | 41.43 | 41.31 | 42.13 | 41.86 | 41.70 | 0.59 | 0.43 | 0.39 | 0.47 |
| 40.65 | 40.41 | 40.33 | 41.15 | 40.86 | 40.77 | 0.50 | 0.45 | 0.44 | 0.46 |
| 39.54 | 39.04 | 38.97 | 39.76 | 39.44 | 39.40 | 0.21 | 0.39 | 0.44 | 0.35 |
| 37.49 | 36.87 | 37.01 | 38.05 | 37.71 | 37.70 | 0.56 | 0.83 | 0.68 | 0.69 |
| **PoznanStreet** | 39.19 | 38.33 | 37.73 | 39.87 | 39.09 | 38.57 | 0.67 | 0.76 | 0.84 | 0.76 |
| 36.36 | 35.90 | 35.06 | 37.94 | 37.34 | 36.61 | 1.58 | 1.43 | 1.55 | 1.52 |
| 34.26 | 33.82 | 32.90 | 35.95 | 35.53 | 34.63 | 1.70 | 1.72 | 1.73 | 1.72 |
| 31.82 | 31.57 | 30.62 | 33.89 | 33.66 | 32.68 | 2.07 | 2.09 | 2.06 | 2.07 |
| **Dancer** | 37.86 | 36.83 | 37.03 | 38.76 | 37.95 | 38.13 | 0.89 | 1.12 | 1.10 | 1.04 |
| 34.78 | 34.12 | 34.27 | 35.89 | 35.34 | 35.54 | 1.11 | 1.21 | 1.27 | 1.20 |
| 32.30 | 31.78 | 32.00 | 33.38 | 32.95 | 33.15 | 1.08 | 1.17 | 1.16 | 1.13 |
| 30.46 | 30.04 | 30.23 | 31.10 | 30.75 | 30.93 | 0.65 | 0.71 | 0.70 | 0.69 |
|  |  |  |  |  |  | **Avg** | 1.23 | 1.27 | 1.16 | 1.22 |

Table 2 shows the average PSNR degradation is about 1.22dB.

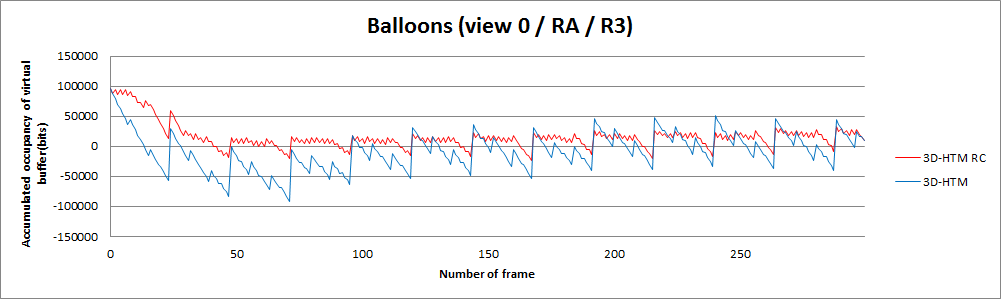
Figures 2 to 8 show the accumulated occupancy of the virtual buffer for the base view of each sequence.



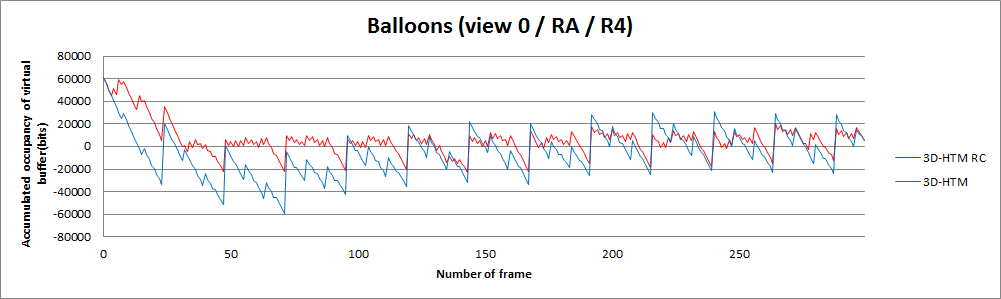
(a)



(b)

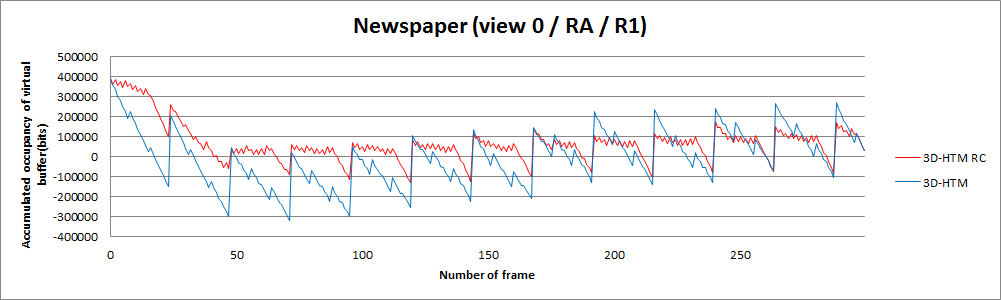


(c)

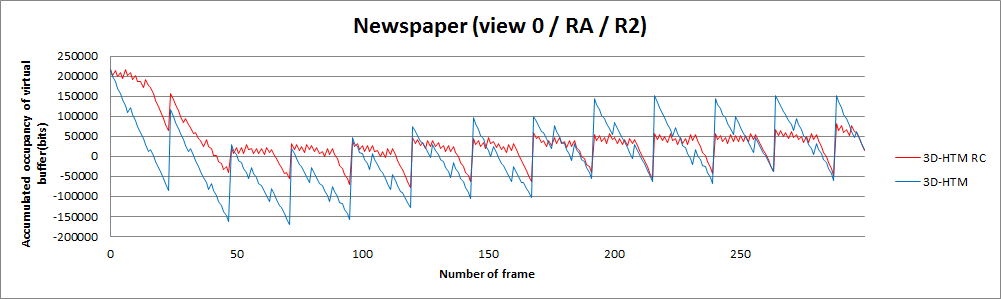


(d)

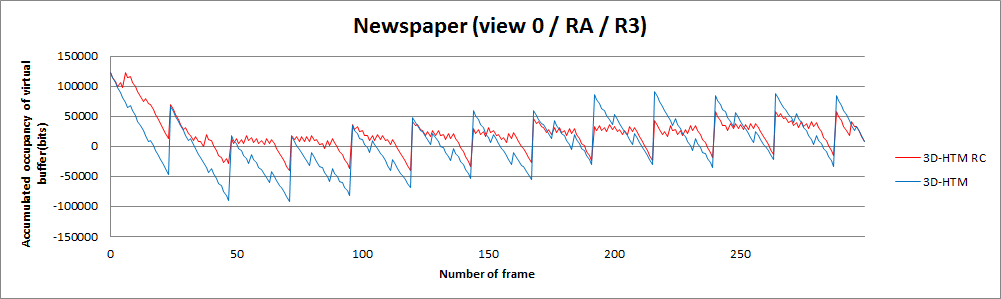
Figure 2. The accumulated occupancy of the virtual buffer for ‘Balloons’ sequence. (a)~(d) : rate1 ~ rate4



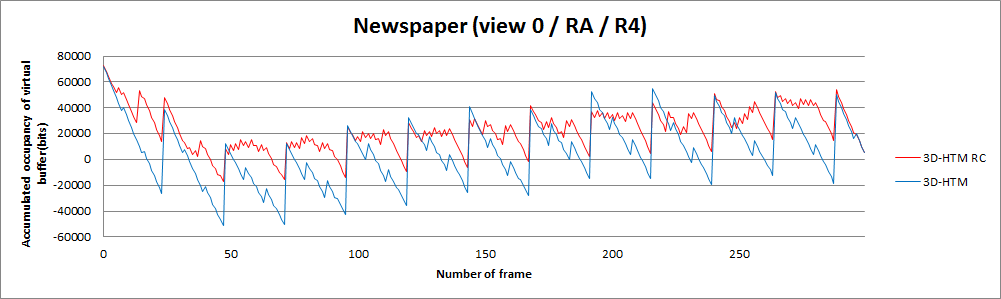
(a)



(b)

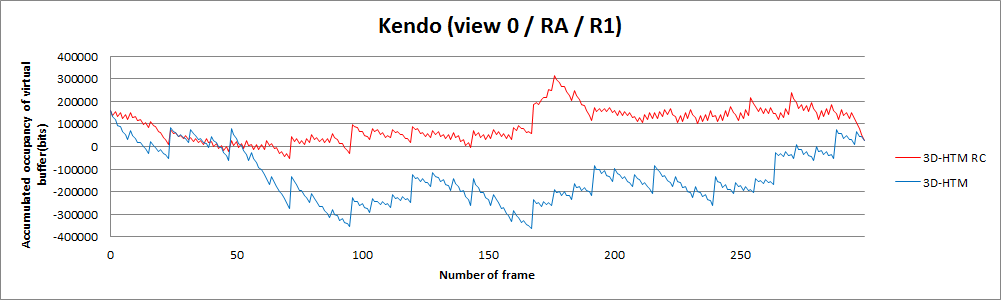


(c)

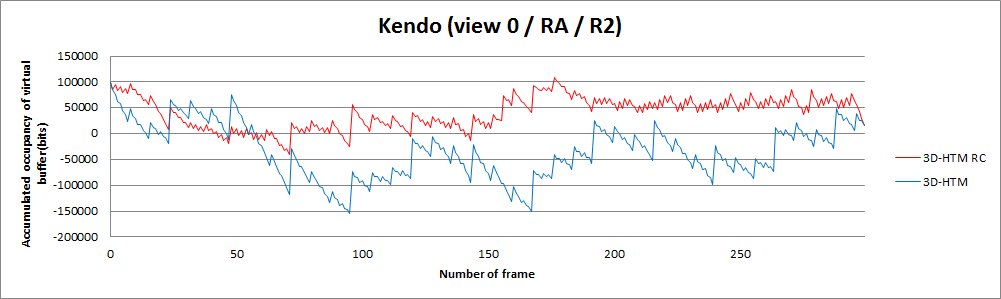


(d)

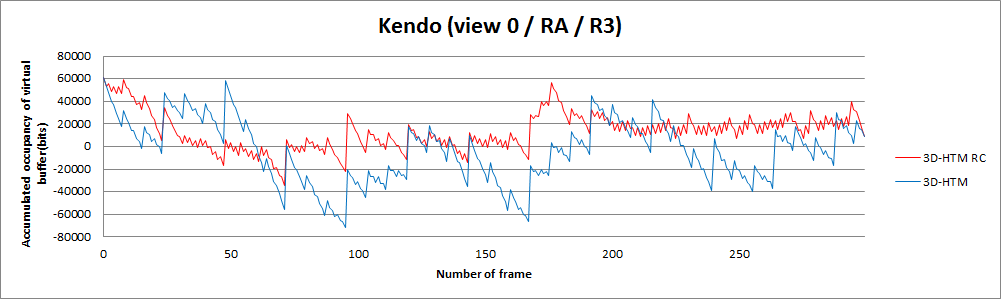
Figure 3. The accumulated occupancy of the virtual buffer for ‘Newspaper’ sequence. (a)~(d) : rate1 ~ rate4



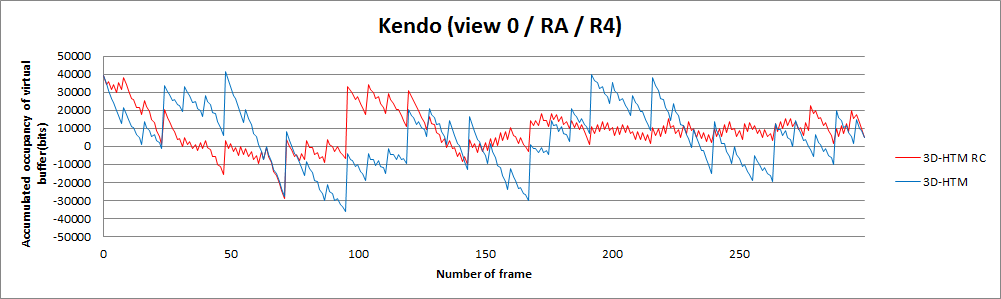
(a)



(b)

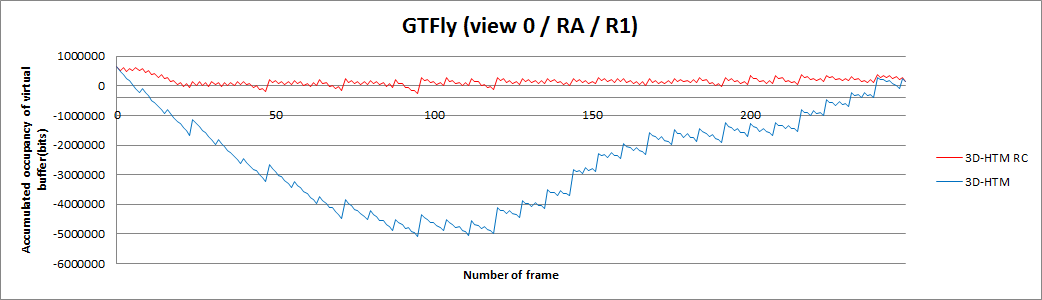


(c)

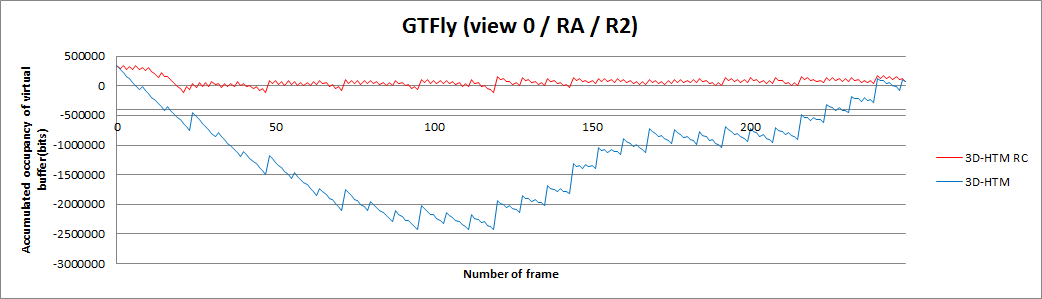


(d)

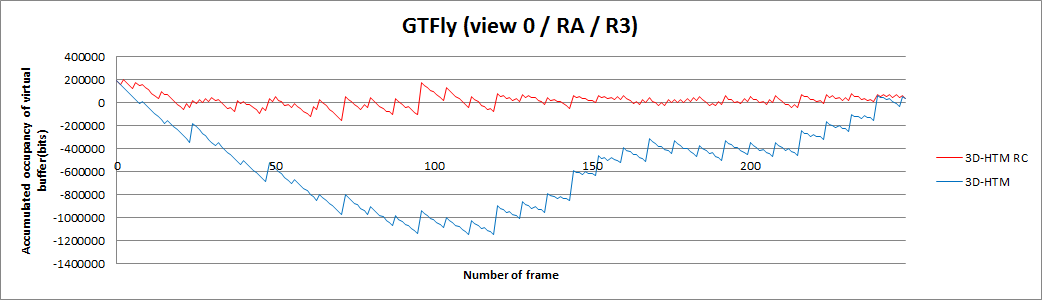
Figure 4. The accumulated occupancy of the virtual buffer for ‘Kendo’ sequence. (a)~(d) : rate1 ~ rate4



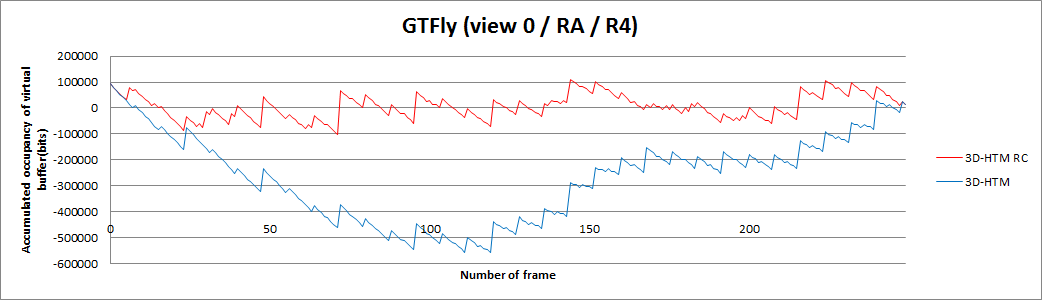
(a)



(b)

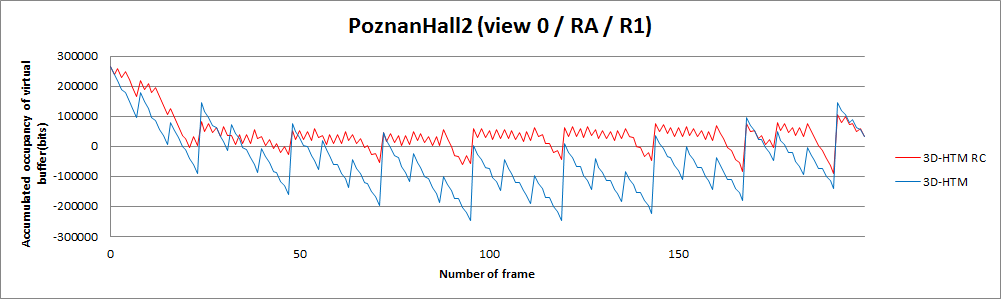


(c)

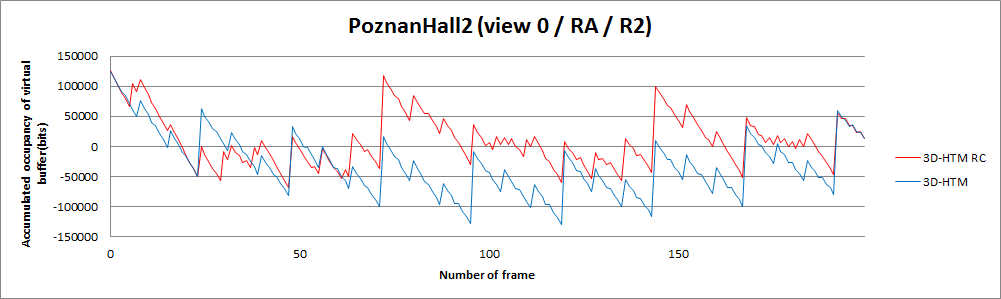


(d)

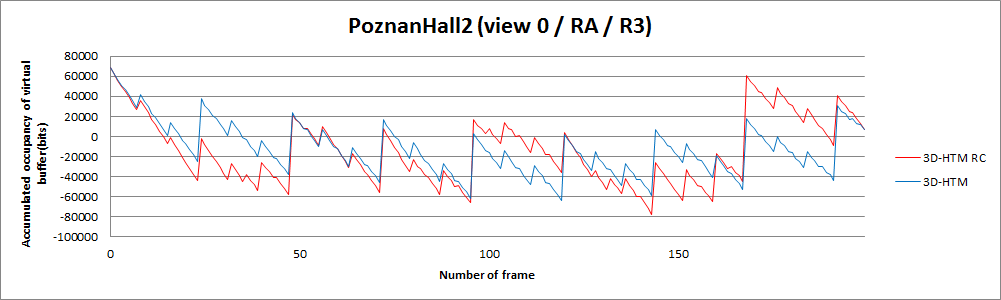
Figure 5. The accumulated occupancy of the virtual buffer for ‘GTFly’ sequence. (a)~(d) : rate1 ~ rate4



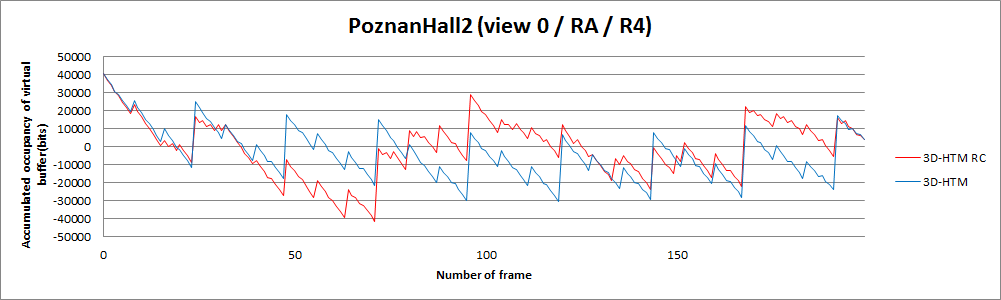
(a)



(b)

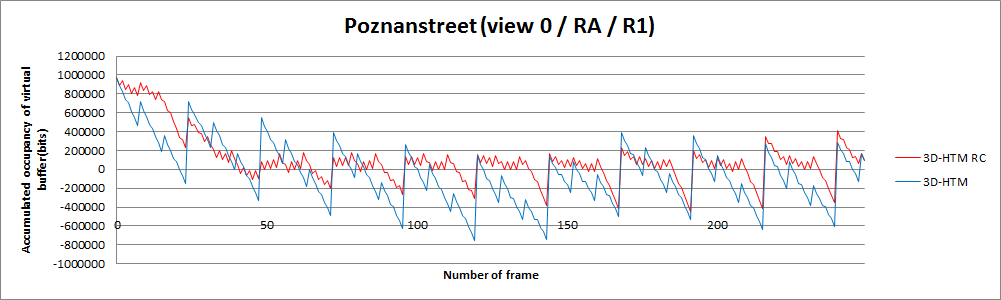


(c)

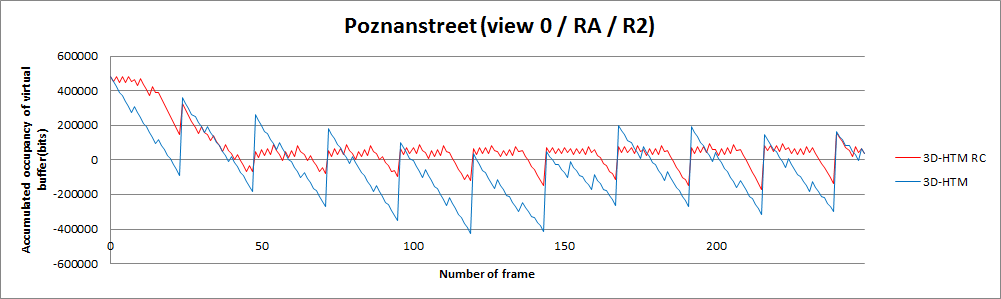


(d)

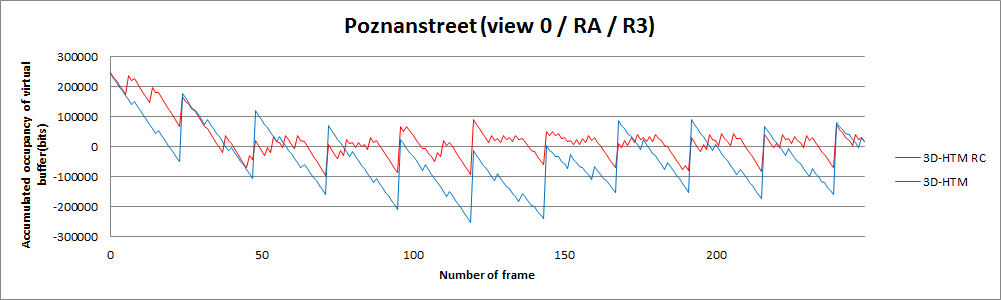
Figure 6. The accumulated occupancy of the virtual buffer for ‘PoznanHall2’ sequence. (a)~(d) : rate1 ~ rate4



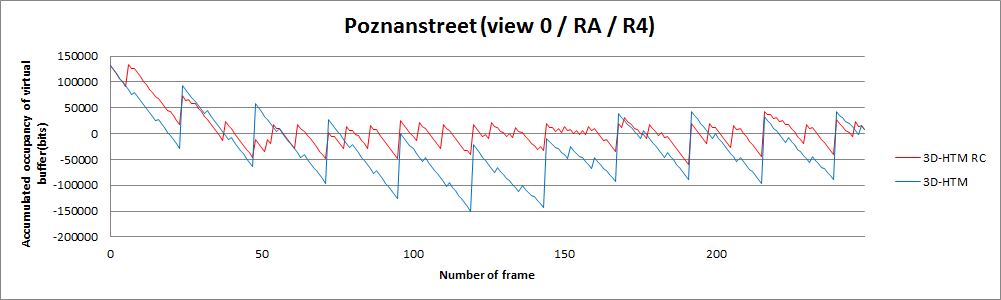
(a)



(b)

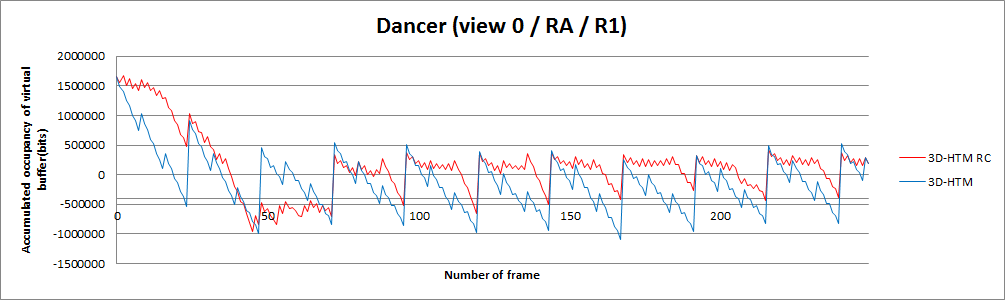


(c)

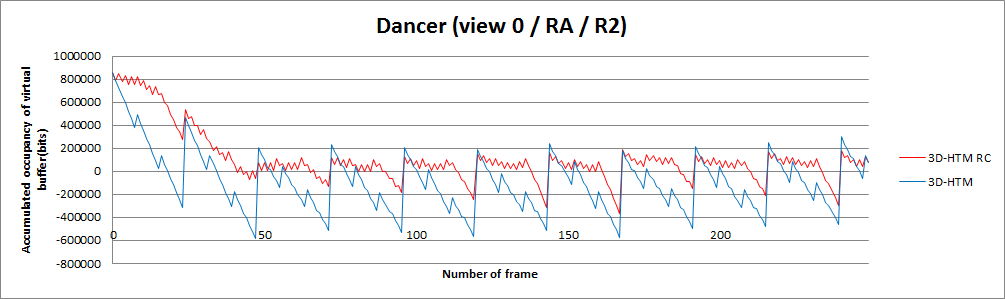


(d)

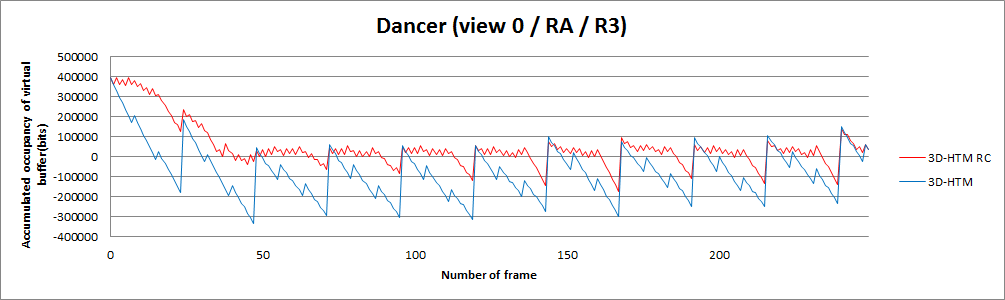
Figure 7. The accumulated occupancy of the virtual buffer for ‘PoznanStreet’ sequence. (a)~(d) : rate1 ~ rate4



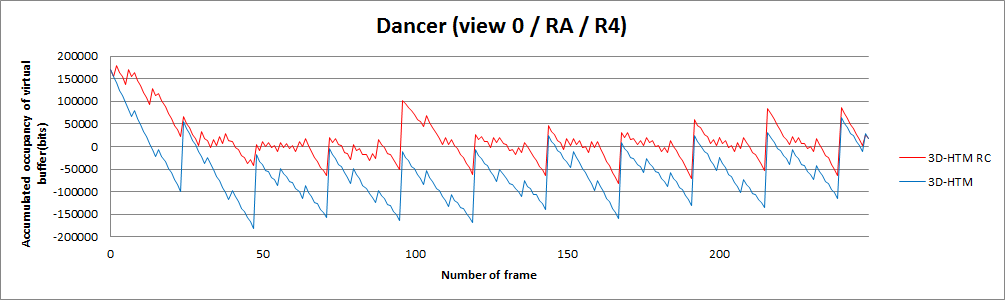
(a)



(b)



(c)



(d)

Figure 8. The accumulated occupancy of the virtual buffer for ‘UndoDancer’ sequence. (a)~(d) : rate1 ~ rate4

According to the above figures, general fluctuation of the accumulated occupancy of the virtual buffer for the base view is reasonable, and the deviation from the target occupancy is much smaller than without rate control. In certain sequences, there is more fluctuation at some bitrates with rate control than without rate control, however, in most cases rate control reduces fluctuation, as expected. The results for the extended views are presented in the additional Excel file included with this document.

# Conclusion

We proposed a rate control scheme for 3D multi-view video coding. The proposed rate control is based on the existing rate control in HEVC. Specifically, the HEVC scheme is applied to the base view directly. In addition, we corrected the initial setting of the QP for the first frame of GOP of extended views. By using this rate control for 3D multi-view video coding, the average bitrate accuracy was 99.2%, with about 1.22dB average PSNR degradation. This is almost the same accuracy for target bitrate but with significantly improved quality of reconstructed video compared to the results presented at the last meeting, where the average PSNR degradation was about 2.1dB. Also, with the newly proposed rate control scheme, most sudden spikes in the number of generated bits were removed, which reduced the fluctuation of the generated bits per frame. In future meetings, we will keep updating the rate control of the base view in the HTM reference software according to the evolution of rate control in HEVC, and we will also keep improving the rate control for extended views.

# Reference

1. W. Lim, H. Choi, J. Nam, H. Jo, D. Sim and I. V. Bajić, “3D-HEVC – Rate control for 3D multi-view video coding,” ISO/IEC JTC1/SC29/WG11 document M25966, Stockholm, Sweden, July 2012
2. H. Choi, J. Nam, J. Yoo, D. Sim, and I. V. Bajić, “Rate control based on unified RQ model for HEVC,” JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCT-VC H0213 (m23088), San José, CA, USA, Feb. 2012
3. H. Choi, J. Nam, J. Yoo, D. Sim, and I. V. Bajić, “Improvement of the rate control based on pixel-based URQ model for HEVC,” JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-I0094 (m24333), Geneva, Switzerland, Apr.-May. 2012.

# Patent rights declaration(s)

**KWU and SFU do not have any current or pending patent rights relating to the technology described in this contribution.**