|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11**  6th Meeting: Torino, IT, 14-22 July, 2011 | Document: JCTVC-F278 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Results about CE1 experiments related to A.1, A.5 and B.x tests** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Guillaume LAROCHE Patrice ONNO Tangi POIRIER  Rue de la Touche Lambert  35510 CESSON-SEVIGNE, France | Email: [guillaume.laroche@crf.canon.fr](mailto:guillaume.laroche@crf.canon.fr)  [patrice.onno@crf.canon.fr](mailto:patrice.onno@crf.canon.fr) |  |
| *Source:* | Canon Research Centre France | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

This contribution presents the results of the experiments conducted in CE1 concerning the motion data storage reduction. The first part of this contribution gives the results about two experiments (A.1 and A.5) regarding the selection of the representative motion data (Category A experiment of CE1). The goal of this experiment is to improve the coding efficiency.

The second part presents all the results regarding the finite-bit representation of motion data corresponding to category B of CE1 experiment. The goal of this experiment is to improve the storage of the motion data.

# Introduction

Core Experiment 1 is about Motion Data Storage Reduction. At the 5th JCTVC meeting some input documents were proposed to modify the Motion Vector storage reduction tool in combination with the temporal MV predictor. No adoption has been done since the current MV Coding and Skip/Merge operations studied in another CE were interacting significantly.

The goal of this Core Experiment 1 is to test the different Motion Data Storage Reduction proposals submitted in the 5th JCTVC meeting in the HM3.0

This contribution relates the results about experiments A.1 and A.5 and all results about the category B of CE1.

# CE1 experiment

## Category A: Selection of the representative motion data

This set of experiments focus on coding efficiency. Several proposals select different MV candidates for motion data storage reduction. This first subset deals with looking at the best Motion Vector representative (MVSR) already tested in the different proposals listed in [1], in combination with the temporal MV predictor (TMVP) location. Related contributions are JCTVC-E059, JCTVC-E092, JCTVC-E142, JCTVC-E147, JCTVC-E211, JCTVC-E221, and JCTVC-E307.

|  |  |  |
| --- | --- | --- |
| **Experiment** | **MVSR (position)** | **Prediction (TMVP position)** |
| Anchor | HM3.0  (TL) | HM3.0  (H or C0) |
| A.1 | BR | HM3.0 |
| A.2 | C0 | HM3.0 |
| A.3 | C3 | HM3.0 |
|  |  |  |
| A.4 | HM3.0 | C0 |
| A.5 | BR | C0 |
| A.6 | C0 | C0 |
| A.7 | C3 | C0 |
|  |  |  |
| A.8 | HM3.0 | C3 |
| A.9 | BR | C3 |
| A.10 | C0 | C3 |
| A.11 | C3 | C3 |
| A.12 | Largest size | HM3.0 |
|  |  |  |
| A.13 | HM3.0 | BR |
| A.14 | BR | BR |
| A.15 | C0 | BR |
| A.16 | C3 | BR |

Table 2: Summary of the experiments for category A.



Figure 1: Description of the positions for MVSR.



Figure 2: Description of the positions for TMVP.

## Results for experiments A.1

In that case, the bottom right (BR) candidate is systematically used for MVSR, and there is no change for the temporal motion vector representative (AMPR), the same representative as in HM3.0 is used.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
|  |  |  | Random access |  |  | Random access LC |  |  |
|  | Y BD-rate | U BD  -rate | V BD-rate | Y BD-rate | U BD  -rate | V BD-rate |  |
|  | Class A | 0.2 | 0.2 | 0.2 | 0.1 | 0.1 | -0.1 |  |
|  | Class B | 0.1 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 |  |
|  | Class C | 0.1 | 0.0 | 0.0 | 0.1 | 0.0 | 0.1 |  |
|  | Class D | 0.3 | 0.3 | 0.2 | 0.3 | 0.1 | 0.0 |  |
|  | Class E |  |  |  |  |  |  |  |
|  | All | 0.2 | 0.1 | 0.1 | 0.1 | 0.1 | 0.0 |  |
|  | Enc Time[%] | 100% | | | 100% | | |  |
|  | Dec Time[%] | 99% | | | 103% | | |  |
|  |  |  |  |  |  |  |  |  |
|  |  | Low delay | | | Low delay LC | | |  |
|  |  | Y BD-rate | U BD  -rate | V BD-rate | Y BD-rate | U BD  -rate | V BD-rate |  |
|  | Class A |  |  |  |  |  |  |  |
|  | Class B | 0.2 | 0.2 | 0.2 | 0.2 | 0.0 | 0.2 |  |
|  | Class C | 0.3 | 0.2 | 0.2 | 0.3 | 0.3 | 0.3 |  |
|  | Class D | 0.3 | 0.0 | 0.6 | 0.4 | 0.3 | 0.1 |  |
|  | Class E | 0.3 | 0.3 | 0.1 | 0.3 | 0.5 | 0.8 |  |
|  | All | 0.3 | 0.2 | 0.3 | 0.3 | 0.2 | 0.3 |  |
|  | Enc Time[%] | 99% | | | 100% | | |  |
|  | Dec Time[%] | 101% | | | 99% | | |  |
|  |  |  |  |  |  |  |  |  |

## Results for experiments A.5

In that case, the bottom right candidate is systematically used for the MSVR case and the temporal representative is C0 as depicted in Figure 2.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
|  |  |  | Random access |  |  | Random access LC |  |  |
|  | Y BD-rate | U BD-rate | V BD-rate | Y BD-rate | U BD-rate | V BD-rate |  |
|  | Class A | 0.5 | 0.6 | 0.5 | 0.6 | 0.6 | 0.6 |  |
|  | Class B | 0.3 | 0.3 | 0.4 | 0.4 | 0.4 | 0.4 |  |
|  | Class C | 0.5 | 0.5 | 0.5 | 0.5 | 0.4 | 0.5 |  |
|  | Class D | 0.4 | 0.3 | 0.3 | 0.3 | 0.2 | 0.2 |  |
|  | Class E |  |  |  |  |  |  |  |
|  | All | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 |  |
|  | Enc Time[%] | 99% | | | 99% | | |  |
|  | Dec Time[%] | 98% | | | 98% | | |  |
|  |  |  |  |  |  |  |  |  |
|  |  | Low delay | | | Low delay LC | | |  |
|  |  | Y BD-rate | U BD-rate | V BD-rate | Y BD-rate | U BD-rate | V BD-rate |  |
|  | Class A |  |  |  |  |  |  |  |
|  | Class B | 0.6 | 0.5 | 0.4 | 0.7 | 0.3 | 0.3 |  |
|  | Class C | 0.6 | 0.6 | 0.6 | 0.7 | 0.5 | 0.3 |  |
|  | Class D | 0.4 | 0.1 | 0.8 | 0.5 | 0.1 | 0.4 |  |
|  | Class E | 0.8 | 0.6 | 0.2 | 0.5 | 0.0 | 0.4 |  |
|  | All | 0.6 | 0.4 | 0.5 | 0.6 | 0.2 | 0.3 |  |
|  | Enc Time[%] | 99% | | | 98% | | |  |
|  | Dec Time[%] | 100% | | | 100% | | |  |
|  |  |  |  |  |  |  |  |  |

## Conclusion for A.1 and A.5 experiments

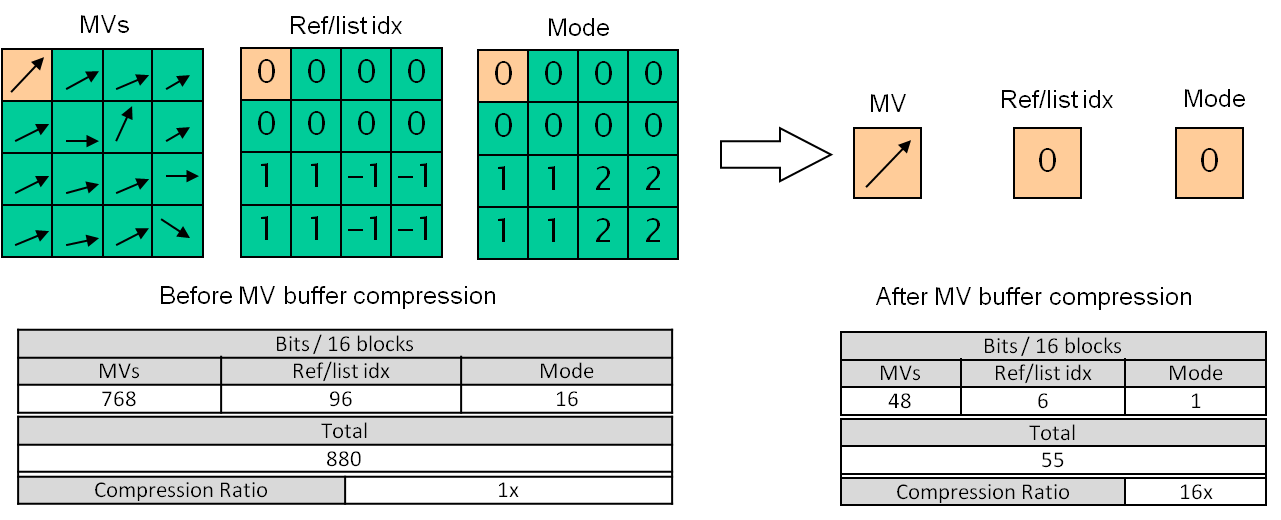
The two experiments tested in A.1 and A.5 show losses compared to HM3.0 anchors.

# Category B: Finite-bit representation and scaling of motion data

This set of experiments aims at reducing the motion data storage while preserving the coding efficiency.

## Current HM3.0 motion data reduction

Figure 3 illustrates the current method for the motion data storage in HM3.0 where only one representative is selected to summarize the motion vectors of 16 4x4 blocks, containing 16 motion vectors, 16 reference index and 16 coding modes per list. Consequently, if we consider that 16 bits are used for each motion vector component and that a maximum of 4 reference frames could be used in each list, 71 bits need to be stored for each block of a B Slice as depicted in Figure 3. Without motion vector data storage reduction, the buffer needed for the derivation of the temporal predictor is approximately 37 Mbits/frame for a 4Kx2K video sequence (37Mbits/frame ≈ ((4096x2048)/(4x4))\*71 bits).



71

64

1024

1136

Figure : HM3.0 motion vector, reference/list index and mode memory.

The number of reference frames is fixed to 4 in each list (L0, L1) in the current testing conditions for HM3.0 and an example of the motion field for a given picture using 4 reference frames per list is illustrated in Figure 4.



Figure : Schematic representation of current HM3.0 reference frame and lists.

## CE1 experiments for category B

Different methods to reduce the motion data storage have been presented during the 5th JCT-VC meeting. The goal of these category B experiments is to evaluate the BDR performance of each motion data reduction method and their combination as presented in Table 1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Experiments** | **Finite-bit MV representation** | | **Reference scaling**  **(E221)** | **Numbers of bits per**  **4x4blocks** |
| **bits per MV component**  **(E142)** | **LSBs removed**  **(E221)** |
| Anchor | 16 | N.A | NO | (16+16)+(16+16)+(2+2)+(1+1) +1 =71 |
| B.1 | 10 | 0 (fixed) | NO | (10+10)+(10+10)+(2+2)+(1+1) +1  =47 |
| B.2 | 0 (Ref only) | (10+10)+ (10+10)+(1+1)+1  =43 |
| B.3 | 1 (One single list) | (10+10)+1  = 21 |
| B.4 | 1 (adaptive) | NO | (10+10)+(10+10)+(2+2)+(1+1) +1  =47 |
| B.5 | 0 (Ref only) | (10+10)+ (10+10)+(1+1)+1  =43 |
| B.6 | 1 (One single list) | (10+10)+1  = 21 |
| B.7 | 8 | 0 (fixed) | NO | (8+8)+(8+8)+(2+2)+(1+1) +1  =39 |
| B.8 | 0 (Ref only) | (8+8)+ (8+8)+(1+1)+1  =35 |
| B.9 | 1 (One single list) | (8+8)+1  = 17 |
| B.10 | 1 (adaptive) | NO | (8+8)+(8+8)+(2+2)+(1+1) +1  =39 |
| B.11 | 0 (Ref only) | (8+8)+ (8+8)+(1+1)+1  =35 |
| B.12 | 1 (One single list) | (8+8)+1  = 17 |

Table : Summary of experiments B of CE1

### Finite-bit Motion vector representation

The goal of this method is to evaluate the impact in terms of coding efficiency when reducing the numbers of bits to store the motions vectors.

The method consists in reducing the number of bits on which each motion vector component is stored. In HM3.0, we assume that the motion vectors are stored on 16 bits. Two experiments have been conducted where this initial number of bits has been reduced to either 10 bits or 8 bits. This experiment is presented in the second column of Table 1.

The reduction of the numbers of bits can be performed on the most significant bit by reducing the motion amplitude or on the least significant bit by reducing the motion vector accuracy. Two different methods have been tested in that core experiments to apply this reduction.

1. In a first method (called “fixed” in Table 1), the reduction of bits is applied systematically on the most significant bits on the motion vectors thus reducing the possible range of each vector component from 16 bits to 10 or 8 bits. This method corresponds to a classical clipping operation of the motion vector components as depicted in Figure 5.

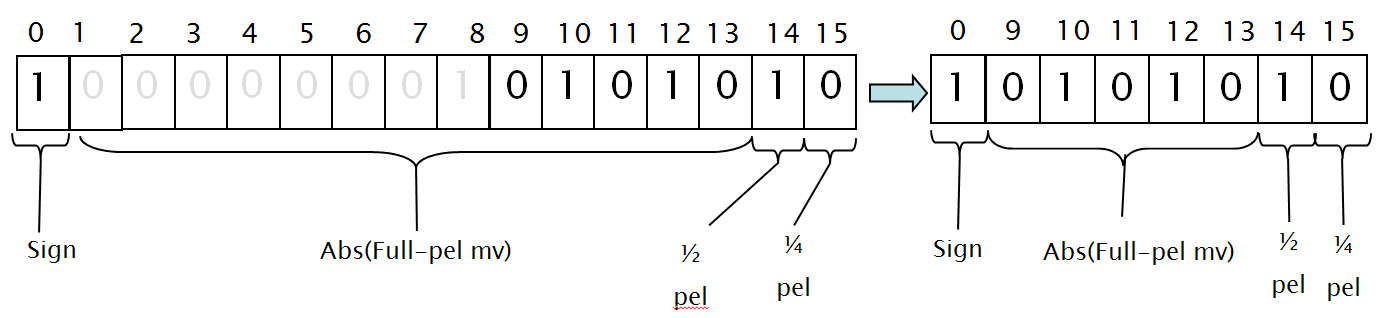


Figure : Motion vector shifting by removing the 8 most significant bits

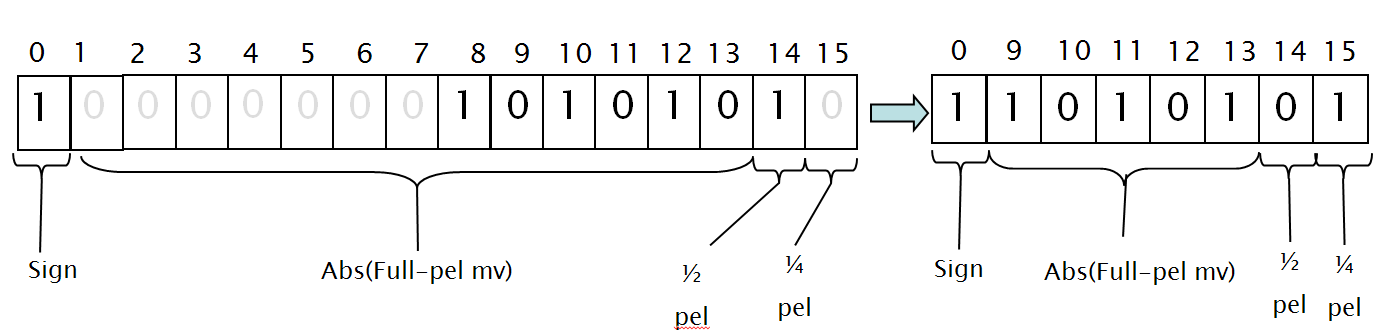


Figure : Adaptive motion vector shifting by removing 7 most significant bits and the least significant bit.

1. In the second method (called “adaptive” in Table 1), the reduction of bits has been applied both on the most significant and the least significant (cf. Figure 6) bits according to the following rule:
   1. Firstly, the least significant bits are removed. The number of bits to remove for the least significant bits is signaled through a table introduced in the Slice Header for each reference frame (up to 4 in HM3.0) and for each list used in the current slice.

|  |  |
| --- | --- |
| List 0 | List 1 |
|  |  |
| ( | ( |
| ( | ( |
| ( | ( |

* + 1. In this core experiment, these numbers are determined as follows at the encoder side: for each reference frame , the POC difference between the current frame and the reference frame in the list is considered and the number is determined according to the formula below. The determination of the number of bits is performed independently on the four references frames of the two lists L0 and L1.
    2. Practically, this consists in removing the ¼ pel accuracy of motion vectors when the distance between the current frame and the reference frame is higher than 1 and remove the ½ pel accuracy when the distance is higher than 4 as showed in the following table.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | | | | | | | |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | … |
| Number of LSB removed | 0 | 1 | 1 | 1 | 1 | 2 | 2 | 2 |

* 1. Then, according to the maximum bits on which the motion vector is represented (8 or 10) the remaining bits to remove correspond to the most significant bits of the motion vectors.

### Scaling in the nearest reference

In addition and in order to further reduce the number of bits to represent the motion data, scaling process to the nearest reference frame has been applied and two configurations have been tested for this method.

#### Two lists, one single reference in each list.

This modification enables to reduce the number of bits for each 4x4 block by scaling the motion vectors of each block in the nearest reference frame of the current frame. In that case, the number of bits (2) allocated to indicate the reference index among four possible references can be avoided. Figure 5 schematically illustrates the results of the scaling process when considering the two lists. Table 2 gives the final number of bits when applying this method where 8 bits are allocated to each motion vector components.



Figure : Schematic representation of the scaling process with one reference per list.

|  |  |  |  |
| --- | --- | --- | --- |
| L0 | | L1 | |
| Mvx | 8 bits | Mvx | 8 bits |
| Mvy | 8 bits | Mvy | 8 bits |
| Ref idx | 2 bits | Ref idx | 2 bits |
| List idx | 1 bit | List idx | 1 bit |
| Mode | | | |
| 1 bit | | | |
|  | | | |
| Total | | | |
| 35 bits | | | |

Table : Number of bits to store motion data information related to each collocated vector when considering the nearest reference frame in each list.

#### One single list and one reference frame.

Moreover, to avoid the storage of the list index and to keep one motion vector per block, in this second method only list L0 is stored. When the motion vector of list L0 is not available, its value is changed by the scaling of the motion vector at the same position in list L1. This process is illustrated in Figure 6. For instance, the 1st top block has two motion vectors. Only the L0 MV is kept, after rescaling. The 2nd top block uses L0 and ref 0, so its MV is stored as is. The 3rd top block uses L1, so its MV is rescaled to refer to ref 0 of L0.



Figure : Schematic representation of the scaling process with one single reference frame

This modification, combined with the previous one described in section 3.2.1, significantly reduces the amount of bits needed for the storage of the motion information of one block. As depicted in Table 3, our proposal removes the information regarding the components of one motion vector, the reference frame indexes buffer and the list indexes buffer. Consequently, no data access is needed for these data.

Finally, when considering the case where only 8 bits are used per motion vector, only 17 bits for each block are stored instead of 71 bits for the current implementation of the motion vector data storage method in HM3.0.

|  |  |  |  |
| --- | --- | --- | --- |
| L0 | | L1 | |
| Mvx | 8 bits | Mvx | 8 bits |
| Mvy | 8 bits | Mvy | 8 bits |
| Ref idx | 2 bits | Ref idx | 2 bits |
| List idx | 1 bit | List idx | 1 bit |
| Mode | | | |
| 1 bit | | | |
|  | | | |
| Total | | | |
| 17 bits | | | |

Table : Number of bits to store motion data information related to each collocated vector when considering the nearest reference frame and one single reference frame.

#### Important remark

These modifications allow simplifying the motion vector temporal derivation process in AMVP. Indeed, only the mode needs to be checked for the derivation process instead of checking the availability of each list and comparing the direction, the reference frame index of 2 motion vectors and selecting the one which cross the current frame and/or the one which has the lower temporal distance.

Consequently, the maximum number of operations needed for the proposed scheme is much lower than the maximum number needed for the current derivation process of the collocated predictor in HM3.0. Moreover, the scaling operations of proposed scheme can be easily paralyzed. Possibly, for the Merge Skip mode no scaling is needed for the temporal predictor.

## Results

All results of CE1 category B are included in the attached Excel sheet and are summarized in Table 4.

Figure 9 shows the result in terms of coding performance when considering the reduction factor of the motion vector data information. This figure plots the data reduction factor versus the average BDR for the 4 configurations. The reduction factor of the different storage methods evaluated in that core experiment can vary. According to the number of bits allocated per motion vector and the use of the proposed scaling process in a single reference, the reduction factor can range from 24 (16 in current HM3.0) to 67 and the maximum loss reported in the set is 0.5%.

Figure 10 and Figure 11 show the BDR measure versus the bit rate when respectively 8 bits and 10 bits are used to represent each motion vector components instead of 16 in the current HM3.0. These two figures show that the “adaptive” mode (B4, B5, B6, B10, B11, B12) provides generally better results than the corresponding “fixed” mode (B1, B2, B3, B7, B8, B9). It is important to note that with the adaptive method, a maximum reduction factor of 67 can be reached while introducing a limited average loss of 0.3% BDR.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Red.**  **factor** | **RAHE** | | | **RALC** | | | **LDHE** | | | **LDLC** | | |
|  |  | **BDR (%)** | **enc time (%)** | **dec time (%)** | **BDR (%)** | **enc time (%)** | **dec time (%)** | **BDR (%)** | **enc time (%)** | **dec time (%)** | **BDR (%)** | **enc time (%)** | **dec time (%)** |
| **HM3.0** | 16x | 0.0 | 100% | 100% | 0.0 | 100% | 100% | 0.0 | 100% | 100% | 0.0 | 100% | 100% |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **B1** | 24x | 0.0 | 100% | 100% | 0.0 | 100% | 97% | 0.0 | 100% | 98% | 0.0 | 100% | 99% |
| **B2** | 26x | 0.1 | 100% | 101% | 0.1 | 101% | 99% | 0.1 | 100% | 99% | 0.2 | 100% | 98% |
| **B3** | 54x | 0.2 | 100% | 100% | 0.2 | 100% | 100% | 0.4 | 99% | 99% | 0.6 | 99% | 99% |
| **B4** | 24x | -0.1 | 99% | 101% | -0.1 | 99% | 99% | 0.2 | 99% | 99% | 0.2 | 100% | 99% |
| **B5** | 26x | 0.0 | 100% | 98% | -0.1 | 100% | 101% | 0.2 | 100% | 97% | 0.2 | 100% | 98% |
| **B6** | 54x | 0.0 | 99% | 101% | 0.0 | 99% | 99% | 0.4 | 99% | 99% | 0.6 | 99% | 99% |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **B7** | 29x | 0.3 | 100% | 101% | 0.4 | 100% | 100% | 0.0 | 99% | 100% | 0.0 | 100% | 100% |
| **B8** | 32x | 0.2 | 100% | 102% | 0.3 | 100% | 101% | 0.1 | 99% | 100% | 0.2 | 100% | 100% |
| **B9** | 67x | 0.5 | 100% | 100% | 0.6 | 100% | 99% | 0.4 | 99% | 99% | 0.6 | 99% | 100% |
| **B10** | 29x | -0.1 | 100% | 99% | -0.1 | 100% | 100% | 0.2 | 100% | 100% | 0.2 | 100% | 101% |
| **B11** | 32x | 0.0 | 100% | 99% | 0.0 | 100% | 100% | 0.2 | 100% | 99% | 0.2 | 100% | 102% |
| **B12** | 67x | 0.1 | 100% | 101% | 0.1 | 99% | 101% | 0.4 | 99% | 99% | 0.6 | 99% | 98% |

Table : Table summarizing all the tests for category B of CE1.

Figure : Results when considering the compression ratio

Figure : Results when considering the compression ratio with a 10 bits representation for each motion component.

Figure : Results when considering the reduction factor with a 8 bits representation for each motion component.

# Conclusion

This document presents the result of two experiments A.1 and A.5 in the category A of CE1. These two experiments do not show any improvement over the current HM3.0 anchors.

Regarding the category B experiments, 12 experiments have been conducted. The results show that B10 and B11 experiments provide the best compromise in terms of reduction factor/BDR loss. The reduction factor of B10 and B11 are respectively 29 and 32 (HM3.0 reduction factor is 16) while having less than 0.1% BDR penalty. So we recommend adopting either B10 or B11 experiment of CE11 for the next HM4.0 WD.

# References

1. J. Jung P. Onno and Y.-W. Huang, “**Description of Core Experiment CE1: Motion Data Storage Reduction**”, JCTVC Contribution JCTVC-E701\_r2, 5th Meeting: Geneva, Switzerland, 16-23 March, 2011.
2. F. Bossen., “**Common conditions and software reference configurations**”, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-E700, 5th Meeting: Geneva, Switzerland, March, 2011

# Patent rights declaration(s)

**Canon Research Centre France may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**