|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  13th Meeting: Incheon, KR, 18–26 Apr. 2013 | Document: JCTVC-M0260 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **SCE3: Results of Test 3.3 on Generalized Residual Prediction** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Xiang Li  Jianle Chen  Krishna Rapaka  Marta Karczewicz  Jani Lainema  Kemal Ugur  Christophe Gisquet  Fabrice Le Léannec  Jonathan Taquet  Edouard François  Guillaume Laroche  Patrice Onno | Tel: Email: | [lxiang@qti.qualcomm.com](mailto:lxiang@qti.qualcomm.com)  [cjianle@qti.qualcomm.com](mailto:cjianle@qti.qualcomm.com)  [krapaka@qti.qualcomm.com](mailto:krapaka@qti.qualcomm.com)  [martak@qti.qualcomm.com](mailto:martak@qti.qualcomm.com)  [jani.lainema@nokia.com](mailto:jani.lainema@nokia.com)  [Kemal.Ugur@nokia.com](mailto:Kemal.Ugur@nokia.com)  [christophe.gisquet@crf.canon.fr](mailto:christophe.gisquet@crf.canon.fr)  [fabrice.leleannec@crf.canon.fr](mailto:fabrice.leleannec@crf.canon.fr)  [edouard.francois@crf.canon.fr](mailto:edouard.francois@crf.canon.fr)  [guillaume.laroche@crf.canon.fr](mailto:guillaume.laroche@crf.canon.fr)  [patrice.onno@crf.canon.fr](mailto:patrice.onno@crf.canon.fr) |
| *Source:* | Qualcomm Incorporated  Nokia  Canon Research Centre France | | |

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

# Abstract

This document reports SHVC CE3 results of test 3.3 on generalized residual prediction (GRP). GRP is a predictive coding tool which utilizes previously coded base layer and enhancement layer pictures to form the prediction for enhancement layer picture. It is reported that with bi-linear interpolation filter, short-tap up-sampling filter, and small block size constraints, the computational complexity (in terms of the numbers of multiplications and additions) and memory access of GRP decoding module in the worst case is similar to or lower than that of motion compensation module in HEVC single layer decoding. At the same time, the average luma BD-rate reduction (the average of RA, LD-P, and LD-B configurations) is reported as 2.8%, 3.2%, 3.5%, and 3.0% for four test cases, respectively.

# Introduction

Generalized residual prediction was proposed in [1][2][3][4] to improve the coding efficiency of SHVC. It uses previously coded base layer and enhancement layer pictures to predict the current enhancement layer picture.

# Technical details

## Framework of generalized residual prediction

The framework of GRP is shown in Figure 1 where the case of uni-directional prediction is illustrated.

Let **B**e and **B**b denote the current block in enhancement layer and its up-sampled collocated base layer block, respectively. Let **P**e0 denote the temporal prediction for the block **B**e obtained by using the motion vector MVe0 (where sub-index 0 refers to reference list 0). Similarly, **P**b0 represents the temporal prediction for the block **B**b obtained by using the same motion vector MVe0 on up-sampled base layer reference picture. Then, the inter predicted residue of the base layer block is obtained as

|  |  |  |
| --- | --- | --- |
|  |  |  |

Considering the temporal prediction **P**e0 for block **B**e, the final uni-directional prediction **P** for block **B**e is

|  |  |  |
| --- | --- | --- |
|  |  |  |

where *w* is a weighting factor, which takes the values 0, 0.5, or 1.

When extending this method to the bi-directional prediction case, the final prediction **P** for **B**e is formed as

|  |  |  |
| --- | --- | --- |
|  |  |  |

where sub-index 1 indicates reference list 1.

Sometimes the up-sampled collocated base layer block shows better quality than enhancement layer block for prediction. In such cases, more weighting shall be given to for higher prediction efficiency. Therefore, an additional weighting mode is proposed as

|  |  |  |
| --- | --- | --- |
|  |  |  |

Since the differential block is the same as that in difference domain inter prediction, the additional weighting mode is called diff-0.5 while the original GRP weighting mode is called res-0.5. In total, there are four GRP weighting modes, i.e., (0, res-0.5, 1, diff-0.5).

Generally, weighting factor indicates the prediction mode of GRP. When *w*=0, GRP reduces to single-layer temporal prediction.



**Figure 1 GRP in uni-directional prediction**

In (2) both block and need interpolation when motion vector points to a sub-pixel position. In fact, (2) can be rewritten to (5),

|  |  |  |
| --- | --- | --- |
|  |  |  |

Considering the motion vectors of block and are always the same, the interpolation can be performed directly on differential signal so that only one interpolation is needed and the computational complexity is reduced. Due to the dynamic range requirement of HEVC interpolation module, the values of the differential signal are clipped to have same 8-bit dynamic range as that of the original signal. When res-0.5 mode is used, the calculation of (5) is actually conducted as to avoid undesired rounding error. In this case, the values in differential block are clipped to [0, 255]. For other non-zero weighting modes, the values in differential block are clipped to [-128, 127].

For bi-directional prediction case, the same simplification is also applied.

## Signaling of GRP weighting mode

GRP coding mode (weighting mode w) is signaled at CU level as a weighting index. The weighting indexes 0, 1, 2, and 3 are used to indicate the weighting modes 0, res-0.5, 1, and diff-0.5, respectively. Truncated unary binarization is applied to the weighting mode before CABAC coding.

## Complexity reduction

### Bi-linear interpolation filter

To reduce computational complexity and memory access, bi-linear interpolation filter is used for GRP block when a MV points to a sub-pixel position, as proposed in [6].

In the software, the bi-linear interpolation filter is on by default. A command line argument “GRPIPF” can be used to enable/disable the filter.

### 4-tap GRP up-sampling filter

A short-tap up-sampling filter may be applied to obtain up-sampled base layer picture/block for GRP. As a trade-off between coding performance and complexity, 4-tap up-sampling/smoothing filter shown in Table 1 is employed for luma and chroma components of GRP mode in this proposal.

Table 1 4-tap up-sampling filter for GRP

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Phase | Filter coefficients | | | |
| 0/16 | 9 | 46 | 9 | 0 |
| 1/16 | 10 | 40 | 14 | 0 |
| 2/16 | 8 | 40 | 17 | -1 |
| 3/16 | 6 | 40 | 19 | -1 |
| 4/16 | 4 | 39 | 22 | -1 |
| 5/16 | 2 | 38 | 25 | -1 |
| 6/16 | 2 | 36 | 27 | -1 |
| 7/16 | 1 | 34 | 30 | -1 |
| 8/16 | 0 | 32 | 32 | 0 |
| 9/16 | -1 | 30 | 34 | 1 |
| 10/16 | -1 | 27 | 36 | 2 |
| 11/16 | -1 | 25 | 38 | 2 |
| 12/16 | -1 | 22 | 39 | 4 |
| 13/16 | -1 | 19 | 40 | 6 |
| 14/16 | -1 | 17 | 40 | 8 |
| 15/16 | 0 | 14 | 40 | 10 |

In the software, the 4-tap up-sampling filter is used by default. A command line argument “GRPUSF” can be used to enable/disable the filter.

### Constraints on block size

To further reduce computational complexity and memory access, the block size of GRP mode is constrained as follows.

* GRP mode can only be applied to the PUs no smaller than 8x8.
* In bi-directional prediction, GRP mode can only be applied to the PUs larger than WxH.

To ease the test, WxH is temporarily indicated in SPS by a variable equal to . In the simulations below, two values of WxH, namely 16x8 and 16x16 and tested.

In the software, no bi-directional prediction size constraint on WxH is specified by default. A command line argument “GRPConsSize” can be used to indicate .

### Disabling of GRP on chroma component

In the 12th JCT-VC meeting, some experts expressed interest on the performance of GRP only on luma component. Therefore, a command line argument “GRPConsFilter” is provided. When (GRPConsFilter & 0x02) is not equal to 0, GRP will only apply to luma component.

### Disabling of GRP 4-tap inter-layer filtering in SNR case

To ease tests, GRP 4-tap inter-layer filtering may be disabled in SNR case. The command line argument “GRPConsFilter” can also be used for this. When (GRPConsFilter & 0x01) is not equal to 0, GRP inter-layer layer filtering in SNR case is disabled.

### Fast GRP mode selection at encoder side

In practice, not all GRP weighting modes need to be checked for each CU. The following fast GRP mode selection methods are used for encoder acceleration.

1. When left, top, and top-right CUs are all larger than the current CU, only one weighting mode is checked.
2. For non-merge PUs at temporal level higher than one, two weighting modes (=0, and res-0.5) are checked at most.
3. For CUs with non-2Nx2N partition type, two weighting modes (=0, and res-0.5) are checked at most.

In the software, a command line argument “FastGRP” is provided. By default, the above methods are all enabled.

## Additional motion estimation at encoder side

In the default implementation, motion estimation is conducted only when (i.e. normal pixel domain) and the resulting motion information is shared by all GRP weighting modes to keep a relatively low encoding complexity. To achieve better coding performance, motion estimation can be performed for each GRP weighting mode. For res-0.5, 1, and diff-0.5, the additional motion estimation is conducted in corresponding differential pictures.

In the software, a command line argument “GRPME” is provided. Please note that this method is disabled by default.

# Test Results and Discussions

In this section, GRP is experimentally verified with four different configurations (defined in Table 2). All tests are based on SHVC common test conditions [5]. Simulation results and complexity information are summarized in Table 3. Please refer to attached excel files for details.

Table 2 Configurations of test cases

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | GRP weighting modes | Bi-linear interpolation | 4-tap GRP up-sampling | GRP block size constraint | No GRP for  chroma | No 4-tap GRP filtering in  SNR | Additional ME |
| Test case 1 | 0, res-0.5, 1, diff-0.5 | Y | N | Uni: 8x8  Bi: 16x8 | Y | Y | N |
| Test case 2 | 0, res-0.5, 1, diff-0.5 | Y | Y | Uni: 8x8  Bi: 16x16 | N | N | N |
| Test case 3 | 0, res-0.5, 1 | Y | Y | Uni: 8x8  Bi: 16x16 | N | N | N |
| Test case 4 | 0, res-0.5, 1, diff-0.5 | Y | Y | Uni: 8x8  Bi: 16x16 | Y | N | N |

Table 3 Simulation results and complexity information

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Case | Conf | Y | U | V | EncT | DecT | 8b/8b | 64b/256b | 64b/512b | Mults | Adds |
| Test case 1 | RA | -1.7% | -5.5% | -6.5% | 119% | 100% | 112% | 111% | 112% | 112% | 112% |
| LD-P | -4.1% | -7.7% | -8.7% | 125% | 99% | 121% | 118% | 119% | 122% | 123% |
| LD-B | -2.7% | -6.7% | -7.6% | 116% | 101% | 119% | 116% | 118% | 120% | 120% |
| **Aver** | **-2.8%** | **-6.6%** | **-7.6%** | **120%** | **100%** | **118%** | **115%** | **116%** | **118%** | **118%** |
| Test case 2 | RA | -2.0% | -3.7% | -3.9% | 119% | 126% | 120% | 120% | 120% | 114% | 111% |
| LD-P | -5.3% | -5.7% | -5.1% | 125% | 128% | 140% | 143% | 143% | 130% | 125% |
| LD-B | -3.1% | -4.6% | -4.5% | 116% | 127% | 130% | 130% | 130% | 121% | 117% |
| **Aver** | **-3.5%** | **-4.7%** | **-4.5%** | **120%** | **127%** | **130%** | **131%** | **131%** | **122%** | **118%** |
| Test case 3 | RA | -2.0% | -3.9% | -4.3% | 114% | 126% | 118% | 119% | 118% | 113% | 111% |
| LD-P | -4.8% | -5.4% | -5.0% | 119% | 128% | 138% | 143% | 143% | 130% | 125% |
| LD-B | -2.9% | -4.7% | -4.7% | 111% | 127% | 127% | 128% | 128% | 120% | 117% |
| **Aver** | **-3.2%** | **-4.6%** | **-4.7%** | **115%** | **127%** | **128%** | **130%** | **130%** | **121%** | **118%** |
| Test case 4 | RA | -1.6% | -4.1% | -5.1% | 120% | 126% | 110% | 109% | 109% | 105% | 104% |
| LD-P | -4.8% | -6.7% | -8.0% | 126% | 127% | 120% | 119% | 120% | 113% | 109% |
| LD-B | -2.6% | -5.1% | -6.0% | 117% | 126% | 114% | 112% | 113% | 108% | 105% |
| **Aver** | **-3.0%** | **-5.3%** | **-6.4%** | **121%** | **126%** | **115%** | **113%** | **114%** | **109%** | **106%** |

Complexity of the proposed tool in the worst case is summarized in Table 4. For all these tests, the computational complexity and memory access of GRP decoding module in worst case is kept similar to or lower than that of motion compensation module in HEVC single layer decoding according to the updated AHG17 template [7].

The number of reference pictures used by GRP is 9 (4 enhancement reference pictures, 4 base reference pictures, and 1 collocated base picture). In contrast, SHM-1.0 only uses 5 (4 enhancement reference pictures and 1 collocated base picture). It looks that GRP increases the number of reference by 80%. In fact, it is not since all these reference pictures are already stored in DPB so that no additional picture buffer is needed, especially when decoding on the fly. Please note that if an enhancement reference does not have collocated base reference, GRP cannot be applied to such a reference. Therefore, GRP does not necessarily lead to larger picture buffer.

The size of lookup table is increased by GRP since additional coefficients of MC interpolation filter and up-sampling/smoothing filter are stored. Compared to SHM-1.0, bilinear MC interpolation filter and 4-tap up-sampling/smoothing filter are newly introduced. In total, the table size is increased by 2x4+2x8+4x16=88 bytes, which is not critical in practice.

Table 4 Complexity of the proposed tool in the worst case

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Test case | Mul | Adds | MemBand (4x2) | MemBand (8x2) | Ref num in pred | Additional picture buffer | Size of lookup table |
| Test case 1 | 110% | 111% | 100% | 98% | 180% | 0% | 113% |
| Test case 2 | 76% | 67% | 108% | 106% | 180% | 0% | 146% |
| Test case 3 | 76% | 67% | 108% | 106% | 180% | 0% | 146% |
| Test case 4 | 56% | 49% | 70% | 68% | 180% | 0% | 146% |

# Conclusions

In this proposal, the performance of generalized residual prediction for SHVC is investigated in different configurations. By utilizing base layer picture information in predicting enhancement layer picture, significant improvement in coding efficiency is achieved. More important, the computational complexity (in terms of the numbers of multiplications and additions) and memory access of GRP decoding module in the worst case is kept similar to or lower than that of motion compensation module of HEVC single layer decoding according to AHG17 template. It is suggested to adopt GRP into SHVC and reference software.

# References

1. X. Li, J. Chen, K. Rapaka, M. Karczewicz, “TE3: Results of Test 4.6.2.1 on Generalized Residual Prediction”, JCTVC-L0078, Geneva, Switzerland, 14–23 Jan. 2013.
2. X. Li, J. Chen, K. Rapaka, M. Karczewicz, “Non-TE3: Extension of Test 4.6.2.1 on Generalized Residual Prediction”, JCTVC-L0190, Geneva, Switzerland, 14–23 Jan. 2013.
3. J. Lainema, K. Ugur, “TE3: Results of test 4.6.3 on base enhanced motion compensated prediction”, JCTVC-L0038, Geneva, Switzerland, 14–23 Jan. 2013.

1. [C. Gisquet](mailto:christophe.gisquet@crf.canon.fr), [F. Le Léannec](mailto:fabrice.leleannec@crf.canon.fr), J. Taquet, [E. François](mailto:edouard.francois@crf.canon.fr), [G. Laroche](mailto:guillaume.laroche@crf.canon.fr), [P. Onno](mailto:patrice.onno@crf.canon.fr), “TE3: Results of test 4.6.1 on the Generalized Residual Inter-Layer Prediction (GRILP)”, JCTVC-L0100, Geneva, Switzerland, 14–23 Jan. 2013.

1. [X. Li](mailto:lxiang@qti.qualcomm.com), [J. Boyce](mailto:jill@vidyo.com), [P. Onno](mailto:patrice.onno@crf.canon.fr), [Y. Ye](mailto:yan.ye@interdigital.com), “Common SHM test conditions and software reference configurations”, JCTVC-L1009, Geneva, Switzerland, 14–23 Jan. 2013.

1. [W. Jang](mailto:wonkap@vidyo.com), [J. Boyce](mailto:jill@vidyo.com), [A. Abbas (Vidyo)](mailto:adeel@vidy.com), [E. Alshina](mailto:elena_a.alshina@samsung.com), [C. Kim](mailto:dionism@samsung.com), “TE3: Results of test 4.4.2 on inter prediction based on difference coding”, JCTVC-L0184, Geneva, Switzerland, 14–23 Jan. 2013.

1. [E. François](mailto:edouard.francois@crf.canon.fr), [A. Tabatabai](mailto:ali.tabatabai@am.sony.com), [E. Alshina](mailto:elena_a.alshina@samsung.com), BoG report: Methodology for evaluating complexity of combined and residual prediction methods in SHVC, JCTVC-L0440, Geneva, Switzerland, 14–23 Jan. 2013.

# Patent rights declaration(s)

**Qualcomm Inc. 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).**

**Nokia 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).**

**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).**