|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  14th Meeting: Vienna, AT, 25 July – 2 Aug. 2013 | Document: JCTVC-N0234 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Low-Complexity Generated Inter-layer Reference for SHVC** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Xiang Li  Jianle Chen  Marta Karczewicz | Tel: Email: | [lxiang@qti.qualcomm.com](mailto:lxiang@qti.qualcomm.com)  [cjianle@qti.qualcomm.com](mailto:cjianle@qti.qualcomm.com)  [martak@qti.qualcomm.com](mailto:martak@qti.qualcomm.com) |
| *Source:* | Qualcomm Incorporated | | |

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

# Abstract

In this proposal, a second inter-layer reference is generated based on previously coded pictures by a GRP (generalized residual prediction) alike method. To reduce computational cost and memory bandwidth requirement, sub-pixel motion compensation interpolation is avoided by rounding motion vectors to the closest integer-pixel positions. It is reported that 0.8%, 1.3% and 0.7% luma BD-rate reduction was obtained on average for RA, LD-B and LD-P cases, respectively.

# Introduction

Generalized residual prediction (GRP) was proposed in [1][2][3][4][5][6] to improve the coding efficiency of IntraBL based SHVC. It uses previously coded base layer and enhancement layer pictures to predict the current enhancement layer picture. In [7][8], the idea was extended to RefIdx based SHVC by generating a new inter-layer reference picture with GRP alike methods. Although good coding performance was obtained by [7][8], the computational complexity and memory bandwidth requirement are high due to the additional motion compensation process introduced during the new inter-layer reference generation. In this document, a low-complexity inter-layer GRP reference generation method is proposed.

# Technical details

It is proposed to generate a new inter-layer reference picture (GRPRef) based on enhancement references, base references, and coded motion field of the collocated base picture. Then insert this generated GRPRef as a second interlayer reference picture, in additional to the resampled base layer picture to code an enhancement layer picture. To reduce computational complexity and memory bandwidth requirement during the GRPRef generation, it is further proposed to round motion vectors to integer-pixel positions so that no motion compensation interpolation is needed for both luma and chroma components.

## Generation of GRPRef picture

Before coding an enhancement picture, GRPRef is generated. The GRPRef is generated by adding the difference signal of enhancement layer reference picture and base layer reference picture onto the resampled collocated base layer picture. To be in line with inter-layer motion mapping, GRPRef is generated by 16x16 blocks.

For each 16x16 block in the current enhancement picture, let denote upsampled (if necessary) collocated base layer block, , represent the enhancement reference and up-sampled (if necessary) base reference indicated by the motion information (reference index) mapped from collocated base layer, the subscript x denotes reference list, respectively. The related block in GRPRef reference is generated as in (1) for uni-directional prediction and (2) for bi-directional prediction.

|  |  |  |
| --- | --- | --- |
|  |  | (1) |
|  |  | (2) |

where indicates the reference block in reference R with rounded motion vector .

To avoid sub-pixel motion compensation interpolation, mapped motion vector and of the 16x16 block are rounded as in (3) for luma component and (4) for chroma components, respectively.

|  |  |  |
| --- | --- | --- |
|  |  | (3) |
|  |  | (4) |

where “&” indicate bitwise operation “and”.

The generated GRPRef shares the same POC as that of the collocated base picture.

## Reference picture list construction

The generated GRPRef is marked as long-term reference and added into the reference list of enhancement layer. Generally, GRPRef picture may be put to any position of a list. In the current implementation, it is put as the last entry of each list by default.

## Combination of GRPRef and other references

As an inter-layer reference picture, the generated GRPRef is marked as long-term reference picture. According to SHVC constraint, motion vector on GRPRef shall be always zero.

In total, there are three possible combinations of inter-picture prediction on GRPRef.

### Uni-directional prediction directly from GRPRef

In this case, the prediction is as

|  |  |  |
| --- | --- | --- |
|  |  | (5) |

(5) actually mimics the GRP mode res-1 in [5].

### Bi-directional prediction with collocated base picture

In this case, the prediction is as

|  |  |  |
| --- | --- | --- |
|  |  | (6) |

(6) actually mimics the GRP mode diff-0.5 in [5].

### Bi-directional prediction with an enhancement reference

In this case, the prediction is as

|  |  |  |
| --- | --- | --- |
|  |  | (7) |

where indicate the enhancement motion on the enhancement reference . When , (7) turns to

|  |  |  |
| --- | --- | --- |
|  |  | (8) |

That is (7) is able to mimic the GRP mode res-0.5 in [5].

In general, all block level GRP modes in [5] can be mimicked by the proposed GRPRef scheme. Therefore, no additional weighting is needed.

# Analysis of memory access

The additional memory access from the proposed method is introduced during the GRPRef generation. Since motion compensation interpolation is avoided by motion vector rounding, only block copy is needed. In the worst case, a 16×16 block in GRPRef is generated by bi-directional prediction. If up-sampled base layer pictures are reused, 5×16×16 luma samples and 5x(16/2)×(16/2) chroma samples need to be accessed to generate a 16×16 block in GRPRef.

# Test Results

In this section, the proposed method is experimentally verified under SHVC common test conditions defined in [9].



On average, 0.8%, 1.3%, and 0.7% luma BD rate reduction was obtained for RA, LD-B and LD-P cases on average, respectively.

Compared block level GRP methods in [1][2][3][4][5][6], the coding performance degradation is due to

* Scaled base layer motion instead of enhancement layer motion is used
* Motion vector is rounded so that the accuracy of motion further degrades
* Signaling cost is increased since GRP weighting is indicated by reference index in the current framework.
* No GRP smoothing filter is used.
* Multiple weighting modes cannot be mimicked in LD-P configurations. (As discussed in section 2.3, B-frame is needed to mimic weighting modes ref-0.5 and diff-0.5)

In general, it is a good tradeoff between coding complexity and coding performance.

# Conclusions

In this proposal, a second inter-layer reference is generated based on previously coded pictures by a GRP alike method. To reduce computational cost and memory bandwidth requirement, sub-pixel motion compensation interpolation is avoided by rounding motion vectors to the closest integer-pixel positions. Simulations show that 0.8%, 1.3%, and 0.7% luma BD rate reduction was obtained for RA, LD-B and LD-P cases on average, respectively.

# 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.
2. X. Li, J. Chen, K. Rapaka, M. Karczewicz, J. Lainema, K. Ugur, C. Gisquet, F. Le Léannec, J. Taquet, E. François, G. Laroche, P. Onno, “SCE3: Results of Test 3.3 on Generalized Residual Prediction”, JCTVC-M0260, Incheon, Korean, 18–26 Apr. 2013.
3. X. Li, J. Chen, K. Rapaka, M. Karczewicz, “Non-SCE3: Simplified Generalized Residual Prediction”, JCTVC-M0275, Incheon, Korean, 18–26 Apr. 2013.
4. A. Aminlou, J. Lainema, K. Ugur, M. Hannuksela, “Non-CE3: Enhanced inter layer reference picture for RefIdx based scalability”, JCTVC-M0155, Incheon, Korean, 18–26 Apr. 2013.
5. Y. He, Y. Ye, “Non-SCE3: ILR enhancement with differential coding for RefIdx framework”, JCTVC-M0155, Incheon, Korean, 18–26 Apr. 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-M1009, Incheon, Korean, 18–26 Apr. 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).**