|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  12th Meeting: Geneva, CH, 14–23 Jan. 2013 | Document: JCTVC-L0440r1 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **BoG report: Methodoly for evaluating complexity of combined and residual prediction methods in SHVC** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Information | | |
| *Author(s) or Contact(s):* | Edouard François Ali Tabatabai  Elena Alshina | Tel: Email: | [edouard.francois@crf.canon.fr](mailto:edouard.francois@crf.canon.fr) [ali.tabatabai@am.sony.com](mailto:ali.tabatabai@am.sony.com)  [elena\_a.alshina@samsung.com](mailto:elena_a.alshina@samsung.com) |
| *Source:* | BoG chairs | | |

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

# Abstract

This contribution reports the outputs from the BoG on Methodoly for evaluating complexity of combined and residual prediction methods in SHVC. Recommendations about the way of evaluating the complexity for the TE3 technologies related to Combined Intra-, Inter- and Inter-Layer-Predictions are provided. These recommendations may also apply to TE4 proposals on Inter-layer Filtering.

# Introduction

In the context of TE3 related to Combined Prediction in SHVC, several proposals showing noticeable coding efficiency gains have been submitted. These proposals mix several prediction signals, related to the intra-, temporal-, or inter-layer predictions (Annex provides some conceptual details). However there is no objective evaluation of the complexity.

The goal of this BoG is to define a methodology to evaluate the complexity in terms of computations number and memory band-width of combined prediction solutions explored in SHVC.

# Prior solutions used for evaluation complexity in JCTVC

In previous JCTVC exploration works related to Motion Compensation Interpolation Filters, the following methodology has been implemented (cf document JCTVC-F764 ‘Report of BoG meeting on CE3 MC interpolation filter’).

Software modules able to evaluate the average of number of operations and memory bandwidth by running decoder have been integrated in the HM used for MC Interpolation Filters CEs.

A related Excel sheet was used to report the following figures for each proposal:

* Average number of add/multi in all bitstreams
* Average memory bandwidth
* Worst case memory bandwidth

# Proposed methodology

A similar methodology is proposed for the evaluation of Combined prediction solutions in SHVC, with a finer granularity in the measurement of complexity figures.

It is suggested to develop similar software modules to measure the number of operations and memory bandwidth at decoder side in the in-coming SHVC Test Model.

The Excel sheet used for SHVC TEs/CEs will be extended to report the related figures. An example Excel template is provided in attachment, with some explanations provided in Annex2.

The following numbers shall be measured:

* Average number of add/multi in all bitstreams
* Average memory bandwidth, related to:
  + Reference EL pictures
  + Reference BL pictures
  + Co-located BL picture
  + All types
* Worst case memory bandwidth, related to:
  + Reference EL pictures
  + Reference BL pictures
  + Co-located BL picture
  + All types
* Measures related to dynamic range (cf F574)
  + Check if Interpolation process calculations within 32 bits integer (as in HEVC Core MC)
  + Check if all intermediate buffers are 16 bits (as in HEVC Core MC)
* When several modes are tested, report (at least) results for each individual mode or for each incremental addition of modes
  + e.g. GRP w=1 / GRP w=0.5 / GRP w=0.5,1
  + 1 mode means 1 additional RD-check

Latest software modules developed in HM5.0 for CE3 of the 7th JCTVC meeting can be used as a starting example (attached to the BoG report).

**BoG discussions based on this initial draft on Complexity Measurement Methodology (2013 Jan 17)**

* It is decided to address also in this BoG meeting the TE4, since the proposed complexity measurement methodoly is also relevant for the evaluation of upsampling filters complexity.
* Current soft module design (used for MC interpolation filters evaluation) has to be generalized for upsampling part
* The software can also be used for the TE related to upsampling filters (TE4)
* Issue raised by one expert: there is difference between CU and PU concepts, in term of complexity penalty: min PU size is 8x4/4x8 while min CU size is 8x8 – memory access is more critical with PU based signaling
* For the Excel template, the minimum blk size of the evaluated mode is needed
* One expert recommends that complexity should also be measured for picture-level upsampling (much less border issues)
* Other experts are in favour of keeping only the on-the-fly block-based approach to measure the complexity
* It is usefull to include complexity analysis for current HEVC MC interpolation filter complexity (EL) and SMuC up-sampling filter complexity, as reference data

**How to handle non-separable filters?**

* Several comments are in favor of measuring in the same way (separable/non-separable) the number of operations; in addition, an indication if the filter is separable or not (additional architecture or not / existing design) must be provided
* In the provided model, whatever coef values are (even 0 or 1), if phase shift is non-zero, it is recommended to always count all the multiplications involved (as it was previous requested for the MC Interpolation Filters case)
* Additional information could be mentioned regarding dyadic or non-dyadic coefs
* One expert mentions that dyadic coefficients filters are actually not beneficial compared to non-dyadic ones if shifts are different for different positions
* The dynamic of coefficients (max, min, sum values) must be mentionned
* One expert will try to provide an example on how to measure complexity (number of operations) in SIMD based designs, as described in JCTVC-G857 (SIMD Analysis of Some Core Spatial Transforms); But it is decided to stay currently with the operations counting method from MC interp filter experiments.
* Suggestion for memory bandwidth measurement: possibility to use several mv fields for anchors, in order to get comparable results; however this is considered as a solution that would significantly overload the evaluation process.
* What about proposals based on the usage on look-up tables instead of filtering ?
* It will be asked to report the additional memory size needs for all proposed technologies
* Indications if filters are fixed or variable have to be provided

**Regarding intra**

Focus only on texture – no related measures on syntax

**Regarding the proposed constraints on Dynamic Range control**

Comment : this constraint can be very easy to fulfill for fixed filters, but more difficult when using adaptive filters

Recommendation for the coding efficiency performance: focus on luma only as a key measure – for chroma only proposals, we will consider similar measures as luma; luma+chroma results can be provided as supplemental data

**Action points**

* Samsung volunteers to provide a first example of an Excel template, for uni- and bi-pred;
* The Excel template will be shared among the BoG chairs; experts who want to share the Excel file can contact the BoG chairs.
* In addition, Samsung volunteers to start implementing the software modules with help from NEC for IntraBL-only approach (non Reference index framework) on top of the SMuC or Test Model to be issued at the end of this meeting.

# Summary and recommendation

It is recommended:

* To implement software modules to measure the number of operations and memory bandwidth at decoder side in the in-coming SHVC Test Model, for each type of signal (EL temporal reference signal, BL temporal reference signal, BL co-located reference signal).
  + When
    - 1 week after Test model is available (for IntraBL-only); implementation in HM5 of such modules is provided in attachment.
  + Who
    - As the software will be used by several TEs (at least TE3, TE4), is it better to have a branch, or is it up to the TEs participants to implement the complexity measurement features?
    - Can be managed within the AHG ‘software SHVC’ activities
    - Samsung volunteers to implement the software modules on top of the Test Model, for the IntraBL-only framework
    - Inter Digital volunteers to do it for the ref-idx framework based on the IntraBL-only design
* To extend the Excel sheet used for SHVC TEs/CEs to report the related figures.
  + When
    - 2 weeks after end of meeting (as usual TE deadline) for the xls template; a preliminary version is provided as a starting point in attachment.
  + Who
    - Samsung volunteers to provide the 1st draft to TEs participants
    - Suggest for deadline of TE responses: 1 week after the release of code
* To ask proponents to provide the following information (mandatory):
  + Average number of add/multi in all bitstreams
  + Average memory bandwidth, related to:
    - Reference EL pictures
    - Reference BL pictures
    - Co-located BL picture
    - All types
  + Worst case memory bandwidth, related to:
    - Reference EL pictures
    - Reference BL pictures
    - Co-located BL picture
    - All types
  + Measures related to dynamic range (cf F574)
    - Check if Interpolation process calculations within 32 bits integer
    - Check if all intermediate buffers are 16 bits
  + When several modes are tested, report (at least) results for each individual mode or for each incremental addition of modes
    - e.g. GRP w=1 / GRP w=0.5 / GRP w=0.5,1
    - 1 mode means 1 additional RD-check
  + Additional information
    - Additional memory size needs (e.g. additional look-up tables)
    - Minimum block size supported
    - Regarding the filters
      * Fixed / Non-fixed
      * Separable / non-separable
      * Dyadic / non-dyadic coefficients
      * Dynamic of filter coefficients (max, min, sum values)
      * Number of taps
      * Number of phases
* Recommendation for the coding efficiency performance: focus on luma only as a key measure – for chroma only proposals, similar measures as luma will be considered; luma+chroma results can be provided as supplemental data

# BoG e-mail list

[edouard.francois@crf.canon.fr](mailto:edouard.francois@crf.canon.fr)  
[ali.tabatabai@am.sony.com](mailto:ali.tabatabai@am.sony.com)   
[elena\_a.alshina@samsung.com](mailto:elena_a.alshina@samsung.com)

[elena.alshina@gmail.com](mailto:elena.alshina@gmail.com)

[jy.park@lge.com](mailto:jy.park@lge.com)

[chulkeun.kim@lge.com](mailto:chulkeun.kim@lge.com)

[asaxena@sta.samsung.com](mailto:asaxena@sta.samsung.com)

[Jani.Lainema@nokia.com](mailto:Jani.Lainema@nokia.com)

[Cheung.Auyeung@am.sony.com](mailto:Cheung.Auyeung@am.sony.com)

[madhukar@ti.com](mailto:madhukar@ti.com)

[Yan.Ye@InterDigital.com](mailto:Yan.Ye@InterDigital.com)

[Jie.Dong@InterDigital.com](mailto:Jie.Dong@InterDigital.com)

[krapaka@qti.qualcomm.com](mailto:krapaka@qti.qualcomm.com)

[Shan.Liu@mediatek.com](mailto:Shan.Liu@mediatek.com)

[cjianle@qti.qualcomm.com](mailto:cjianle@qti.qualcomm.com)

[patrice.onno@crf.canon.fr](mailto:patrice.onno@crf.canon.fr)

[kenji.kondo@jp.sony.com](mailto:kenji.kondo@jp.sony.com)

[mei.guo@mediatek.com](mailto:mei.guo@mediatek.com)

**ANNEX 1 – examples of basic combined prediction approaches**

# 2nd order residual prediction approaches for Intra prediction

This section describes several basic designs of Combined Prediction mixing Residual prediction (between the EL and BL reference prediction, or DC/gradient based EL prediction sample modification.

## Intra DC Correction (JCTVC-L0036)

DC correction is performed for all intra PUs which have a size of 8x8 or larger and are not IntraBL coded. This mode adds a constant value to each sample of the enhancement layer block so that the DC of the enhancement layer prediction equals to the DC of the co-located base layer block. In other words, a dc\_delta is calculated by taking the difference between enhancement layer block’s average sample value and corresponding upsampled base layer block’s average sample value, and added to all the samples of the enhancement layer block:

P(x,y) = P’(x,y) + dc\_delta

where:

P(x,y) denotes the sample to be predicted in the enhancement layer picture.

P’(x,y) denotes a predicted enhancement layer sample after HEVC intra prediction process

dc\_delta denotes the difference of average sample values in P’(x,y) and corresponding upsampled base layer block.

## Gradient based Intra Prediction (JCTVC-L0037)

For directly horizontal and vertical intra prediction modes, the prediction is enhanced by measuring how the base layer pixel values change in the corresponding direction and applying the generated differential values when reconstructing prediction samples for the enhancement layer video. This process is illustrated in Figure 1 below for a 4x4 block using vertical prediction.

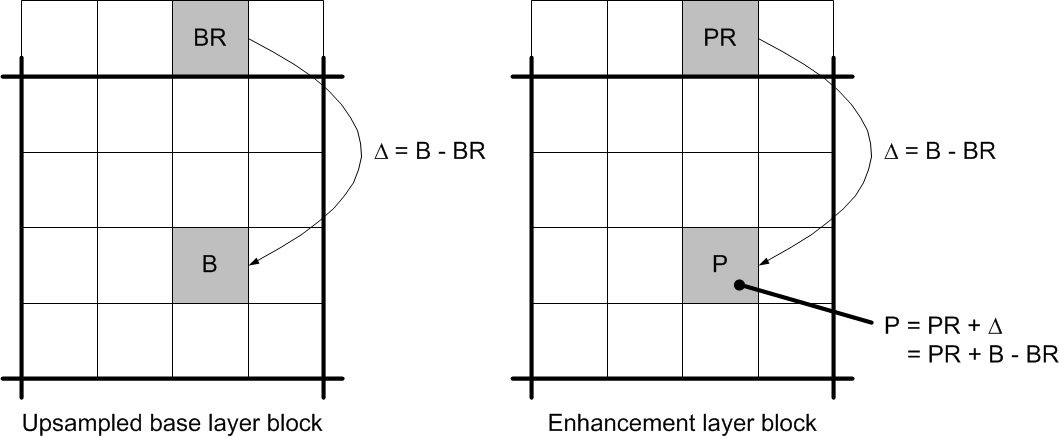


Figure 1: Gradient based enhancement of intra prediction for vertical prediction

## Intra prediction based on difference picture (JCTVC-L0135, JCTVC-L0183, JCTVC-L0222)

This mode corresponds to the Diff INTRA mode, applying intra prediction on the residual inter-layer reconstructed signal. The difference domain Prediction mode is indicated at CU level.



Figure 2: Diff INTRA mode

### Additional changes based on Diff INTRA (JCTVC-L0222)

Planar mode: After intra prediction is performed to get the difference prediction signal, the bottom-right portion of the difference prediction for the planar mode is set to zero (that is each pixel(x, y) satisfying the condition (x + y) >= N-1, where N is the width of the current block is set to zero).

MDIS: HEVC mode dependent intra smoothing (MDIS) process is disabled in difference domain intra prediction mode.

## Filling Unavailable Reference Samples (JCTVC-L0099)

Current padding process is as shown in figure 1. Padding is performed when at least one sample is unavailable for intra prediction. Starting from A, padding is performed by copying its previous pixel (Ppre) to the current (Pcur). If A is not available, set A equal to the first available one (Pnext). If all samples are unavailable, (1 << ( BitDepthY ‑1)) is substituted for the values of all samples. In JCTVC-K0033 [1], filling unavailable reference samples in intra prediction at enhancement layer is proposed for HEVC scalable extension and these techniques were decided to be investigated on top of SMuC-0.1.1 as a part of Tool Experiment 3 (TE3).



Figure 3: normal HEVC padding process for intra reference samples

# 2nd order residual prediction approaches for Inter prediction

This section describes several basic designs of Combined Prediction mixing Residual prediction (between the EL and BL reference prediction, or between the BL reference and BL current predictions) with current upsampled Base layer prediction or temporal EL prediction.

In the sequel, we define as

* PREDEL the prediction signal of the enhancement layer
* REFEL and REFBL the temporal reference signals in the Enhancement and Base layers
* RECBL the base layer reconstructed signal corresponding to the current enhancement layer block
* MVEL the EL motion vector selected
* UPx{.} the up-sampling operator x
* MCx[I,MV] the motion compensation operator x of the current block using I as reference picture and MV for the motion vector.

## Mixed Inter-Layer and Temporal EL prediction

The mixed inter-layer and temporal EL prediction mode consists in mixing the upsampled co-located BL signal with the temporal prediction signal. The prediction signal is derived as follows (shown for the mono-prediction case, the concept is easily generalized for the bi-prediction case):

PREDEL = ( **UP1**{RECBL} + **MC1**[ REFEL , MVEL ] ) / 2



Figure 4: Schematic principle of the Mixed Inter-Layer / Temporal pred mode.

## Diff Inter Mode

The Diff INTER mode consists in applying a second order prediction, using the residual between the enhancement layer temporal reference block and the corresponding upsampled Base Layer temporal reference block using the EL motion vectors. This residual is added to the upsampled base layer block co-located with the current EL block. The prediction signal is derived as follows (shown for the mono-prediction case, the concept is easily generalized for the bi-prediction case):

PREDEL = **UP1**{RECBL} + **MC1**[ ( REFEL - **UP2**{REFBL} ) , MVEL ]



Figure 5: Schematic principle of the Diff INTER mode.

## Generalized Residual Inter-Layer Prediction (GRP)

The GRP mode consists in applying a second order prediction, using the Base Layer temporal residual derived from the scaled EL motion vectors. The prediction signal is derived as follows (shown for the mono-prediction case, the concept is easily generalized for the bi-prediction case):

PREDEL = **MC1**[REFEL, MVEL] + **UP1**{RECBL} – **MC2**[ **UP2**{REFBL} , MVEL ]

The principle of the GRP mode can be summarized by the following Figure 6.



Figure 6: Schematic principle of the GRP mode.

**ANNEX 2 – Memory access evaluation**

This is an explanation of the formulas used in the attached Excel template for memory access worst case evaluation, recommended to be used in SHVC CE3 and CE4.

Notations:

W - width of interpolated block

H - height of interpolated block

w - width of "memory pattern"

h - height of "memory pattern"

TL is filter length for Luma

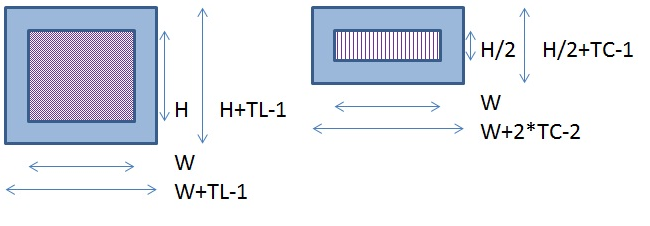
TC is filter length for Chroma

S\_EL – Resolution of layer where we are doing prediction

S\_BL – Resolution of layer where we are referencing

S = S\_BL/S\_EL (1/2 in case ×2 and 2/3 in case ×1.5 and 1 in SNR and MC cases)

1. Luma (Y) block is W × H rectangle
2. Chroma is W × (H/2) rectangle (Cb&Cr are interleaved)
3. In a worst case (2D interpolation) block extension in both direction is needed



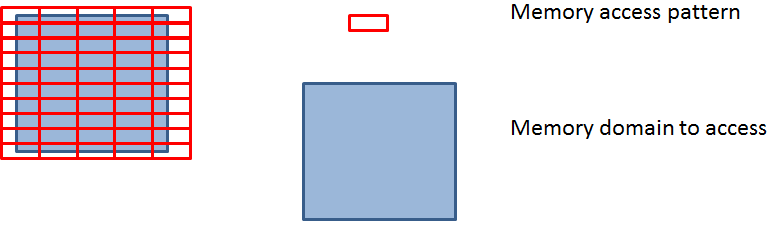
Luma block extended for interpolation. Chroma block extended for interpolation.

1. If predicted block and reference block belong to the layer with different spatial resolution then scaling is needed:

(W+TL-1)×(H+TL-1) 🡪 (W+TL-1)×(H+TL-1) /S/S

We can assume S\_BL = S\_EL (🡪S=1) in case of MC and SNR scalability.

1. We are reading samples from the memory not sample by sample but using w×h patterns. Left-upper corner of memory domain to access and pattern don’t match in the worst case:



1. Resulting size of the memory to access is

(CEILING((W +TL-1)/S**+** w - 1, w))\*(CEILING((H +TL-1)/S**+** h - 1 , h)) - for Luma

and

(CEILING((W +2TC-2)/S**+** w - 1, w))\*(CEILING((H/2 +TC-1)/S**+** h - 1 , h)) - for Chroma

**Dynamic Range analysis**

This is an explanation of the formulas used in the attached Excel template for dynamic range worst case evaluation, recommended to be used in SHVC CE3 and CE4. It assumes that filtering is a 2 stages process.

1st stage

2nd stage

Input parameters for Dynamic range estimation:

SumNeg – sum of negative coefficients of your filter (for {-1, 4,-11,40, 40, -11, 4, -1} SumNeg =-24), can be different for each stage

SumPos – sum of positive coefficients of your filter (for {-1, 4,-11,40, 40, -11, 4, -1} SumPos =88) ), can be different for each stage

MinIn, MaxIn – Min and Max for Input (0 and 255 in most cases)

shift – de-scaling stage (can be different for each stage)

Bit-depth of accumulated sum Reg, and temp buffers are evaluated automatically.

Basic idea embedded into formulas is:

RegMin = SumPos\* MinIn + SumNeg \*MaxIn

RegMax = SumPos\* MaxIn + SumNeg \*MinIn

*If your filter is 1 stage 2D non-separable filter then refer to the example in a template*