|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  21st Meeting: Warsaw, PL, 19–26 June 2015 | Document: JCTVC-U0142 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Inter/Intra Block Copy Unification: Comments and Observations** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | A.M. Tourapis, Y. Su, D. Singer  1 Infinite Loop Cupertino, CA 95014 USA | Tel: Email: | +1-408-228-7983 atourapis@apple.com |
| *Source:* | Apple Inc. | | |

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

# Abstract

At the 20th JCT-VC meeting it was decided to unify the inter prediction and intra block copy processes. This was done, given the strong similarities of the two processes, in an attempt to simplify the specification and the implementation of the intra block copy method, and to allow the use of additional prediction mechanisms, already supported for inter prediction, with the intra block copy framework. This included, for example, bi-prediction and weighted prediction. This contribution makes some additional observations with regards to this unification framework and provides some suggestions on how to better unify the two processes. Intra Block Copy displacement vector limits are also discussed.

# Introduction

A major milestone was reached at the 20th JCT-VC meeting with the unification of the inter prediction and intra block copy (IBC) processes for Screen Content Coding [1]. The IBC process, which is essentially a subset of Fractal Image compression first proposed by Michael Barnsley [2][3], can enable considerably improved compression, especially on intra only images, by allowing prediction of entire blocks from previously decoded areas in the same picture. This process, even though it does not have the full flexibility of fractal compression[[1]](#footnote-1), strongly resembles inter prediction since it essentially requires displacement vectors to be signalled that point to the area to be copied. Previously, IBC partitions required the signalling of new syntax elements that indicate the mode and displacement parameters to be used. In addition, special handling of such partitions had to be performed when performing operations such as prediction and deblocking, among others. With the unification process, however, handling of this mode was considerably simplified.

More specifically, IBC prediction is now performed by introducing a new “dynamic” reference picture inside the Decoder Picture Buffer. This picture, if available, is updated after encoding/decoding a particular region in the current picture, e.g. a CU or LCU depending on the prediction restrictions that one may wish to impose to IBC, with new prediction samples. Only the available samples are allowed for reference purposes, while some restrictions are imposed to assist operations such as wavefront processing. By adding the current picture in the DPB one essentially has the ability to use all tools available for inter prediction including, picture list modification, weighted prediction, and bi-prediction among others. However, the current text still restricts the use of bi-prediction from only the current picture, as well as the use of subpixel prediction from current picture samples.

# Subpixel Support and Weighted Prediction

Extending support of IBC for bi-prediction was already discussed in JCTVC-U0079 [4]. There are also several contributions at this meeting discussing how to best unify motion vector resolution for both IBC and Inter prediction. However, most of them seem to still desire to constrain IBC to integer motion vectors only. We are of the opinion that if a true unification process is desired, then IBC should *also be allowed to use subpixel precision as well*. Subpixel precision can provide additional coding benefits, as suggested also in [5], whereas it can simplify some implementations since no special handling to distinguish between IBC and inter modes is required.

Allowing subpixel precision for IBC enables us to also better unify handling of IBC and inter prediction in other processes of the HEVC, such as the deblocking process, and ends up simplifying the text considerably. The only restrictions that are likely needed, and as also [5] suggests, include *disallowing the use of any vectors, subpixel or otherwise, that may require samples that are considered unavailable*. That implies also subpixel motion vectors at slice/tile boundaries that may require unavailable samples for the interpolation process. Although it might be possible to allow padding of samples, i.e. at image boundaries, such *padding should likely not be allowed internally within the picture* (e.g. at slice/tile boundaries or in the area not allowed for prediction due to wavefront processing). In general, it should be desirable that the use of IBC is completely transparent to the decoder, and any such restrictions are mainly imposed on the encoder as conformance requirements.

Even though the unification process was introduced to simplify the implementation and handling of the IBC mode, we have observed that several new contributions discuss adding several new restrictions that appear to impact syntax elements and/or semantics in relationship to this mode, which we think might be unnecessary. For example we have seen requests to alter the behaviour of the list modification process to accommodate for the availability of the current picture as a reference, disabling weighted prediction for such pictures and so on. We believe that such restrictions are likely unnecessary and are not in spirit with the unification process. As is, and assuming the current picture may be available in both or either list in a B slice, an encoder now has the full flexibility to decide which picture to place in which list. When encoding a B slice, the encoder may decide to use as references only the current picture in both lists, while using different weighted prediction parameters with each instance, or may select to use only temporal references in one list and the current picture reference or references in the other list. The assignment could be based on which one is considered as being of higher priority than the other, for example. Such considerations could also have tremendous benefits for error resiliency/concealment applications, especially in combination with weighted prediction parameters [6][7].

# IBC Displacement Vector Limitations

One of the concerns about the IBC mode is that it may result in an increase in the memory and bandwidth requirements of a video encoder and decoder (especially) system. This is because now the current picture may have to be stored, at least temporarily, in two modes, with and without post-filtering[[2]](#footnote-2). It has been suggested that imposing certain constraints onto the displacement vectors for IBC might be desirable in reducing such penalties. From our analysis, however, we think that this is not really a big issue and one should be able to easily implement with little if any penalty, in both software and hardware, video decoders that can handle this mode while using the entire search area.

Nevertheless, if displacement vector constraints are still desired for IBC, we would like to propose considering the use of a maximum vertical displacement vector, as well as either a horizontal displacement vector or a limit on the N past CTUs that a system may use for prediction. The N past CTUs constrain may be preferable than the horizontal displacement vector in some implementations, since it can better exploit caching. We believe that the vertical limit is necessary in that case to also avoid issues with the use of multiple tiles. Such limitations, if introduced, could be level dependent and should be complementary to, and likely smaller than, the motion vector limits already specified in the Video Usability Information (VUI) syntax of HEVC. In fact, since we believe that the displacement vector limits should be different than the limits imposed for inter vectors, we would also recommend the addition of two additional parameters in the VUI elements that specify the max displacement vectors, and/or the number of past CTUs that may need to be stored for accessing IBC prediction data. These elements could be named as, log2\_max\_ibc\_dv\_length\_vertical, log2\_max\_ibc\_dv\_length\_horizontal, and/or max\_ctu\_no\_ibc\_buffering, for example. We will then also need to specify that:

log2\_max\_ibc\_dv\_length\_vertical <= log2\_max\_mv\_length\_vertical and  
log2\_max\_ibc\_dv\_length\_horizontal<= log2\_max\_mv\_length\_horizontal,

and indicate that **log2\_max\_ibc\_dv\_length\_horizontal** and **log2\_max\_ibc\_dv\_length\_vertical** indicate the maximum absolute value of a decoded horizontal and vertical motion vector component, respectively, in quarter luma sample units, for all reference pictures used for prediction that correspond to the current picture, as dictated by the use of the curr\_pic\_as\_ref\_enabled\_flag parameter.

# Conclusion

This contribution proposes, under the new unification IBC/inter prediction scheme, to allow the use of subpixel precision also for the current picture references. It is also suggested that any restrictions or limitations that would diverge from the principles of unification, unless there is good justification, should likely be avoided. However, additional displacement vector limitations that only apply to references pointing to the current picture may be desired. It is proposed that if those are introduced, these should likely be either level dependent, or introduced within the VUI syntax of the HEVC specification.

# References

1. JCTVC-T1005, “High Efficiency Video Coding (HEVC) Screen Content Coding: Draft 3,” 20th Meeting: Geneva, CH, 10 February – 17 February 2015.
2. M.F. Barnsley, “Fractals Everywhere”, Academic Press, 1988,1992
3. M.F.Barnsley and L.P.Hyrd, “Fractal Image Compression, AK Peters, Ltd., Welessley, 1993
4. K. Rapaka, V. Seregin, R. Joshi, M. Karczewicz, Y.-K. Wang, A. Tourapis, Y. Su, and D. Singer, “HLS : On Intra block copy signaling control”, JCTVC-U0079, 21st JCT-VC Meeting, Warsaw, PL, June 2015
5. K. Rapaka, V. Seregin, R. Joshi, and M. Karczewicz, “Intra block copy chroma interpolation for Non-444”, JCTVC-U0080, 21st JCT-VC Meeting, Warsaw, PL, June 2015
6. Y. Liu and R. Kurceren, “H.264/AVC Error Resilient Video Streaming Using Leaky Prediction”, in Proceedings of Visual Communications and Image Processing 2007 (VCIP’07), Vol. 6508, 2007
7. Y. Liu, R. Kurceren, and D. Mukherjee, “Rate-distortion analysis of weighted prediction for error resilience,” in Proceedings of the 2008 IEEE International Conference on Image Processing (ICIP’07), pp. 2032-5, San Diego, CA, Oct. 2008

# Patent rights declaration(s)

**Apple 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).**

1. E.g. no support for zoom, rotation, iterative refinements, and until recently illumination compensation etc. [↑](#footnote-ref-1)
2. This assumes that the post-filtering processes, i.e. deblocking and SAO, are immediately performed when all related samples are available. There may still be some systems that may wish to perform these processes only after the entire slice/tile/picture has been decoded. [↑](#footnote-ref-2)