|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11**  5th Meeting: Geneva, 16-23 March, 2011 | Document: JCTVC-E241 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Low complexity bi-predictive interpolation with 6-tap DCT-IF filter** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Kemal Ugur Jani Lainema | Tel: Email: | +358 50 4860857 kemal.ugur@nokia.com |
| *Source:* | Nokia | | |

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

# Abstract

This contribution proposes to use a 6-tap DCT-IF filter for interpolation of bi-prediction samples instead of the current 8-tap DCT-IF filter for the low complexity configuration. It is reported that using a filter with fewer taps, the worst-case complexity of the current interpolation filter in HM is reduced. The coding efficiency impact of the proposal is 0.5% on average for the low-complexity cases. However, it is argued that most of the coding efficiency loss is due to single low-resolution BQSquare sequence and for high-resolution Class-A,B, E sequences, the proposal reportedly improves the coding efficiency of low complexity cases by 0.8% on average, while reducing the complexity.

# Introduction

At the Daegu meeting, it was decided to unify the low complexity and high efficiency interpolation filters and an 8-tap DCT-IF filter was chosen for both configurations [1]. This decision improved the coding efficiency for low complexity configuration, however came with a significant increase in complexity.

According the complexity estimates used in CE3 in Daegu meeting, the average complexity increased by around 2x and worst-case interpolation complexity increased by around 3x compared to the low complexity configuration prior to Daegu meeting (as measured in terms of operations / pixel). The increase in worst-case complexity is especially problematic, as many implementations the worst-case complexity has a direct impact on the decoder’s capability.

In order to reduce the interpolation complexity, it is proposed to use a different interpolation filter for bi-prediction that has fewer taps than the filter used for uni-prediction. The main motivation for this proposal is to reduce the worst-case interpolation complexity as the worst-case happens when bi-prediction is used. In addition to reducing the computational complexity, this proposal reduces the worst-case memory bandwidth in motion compensation.

# Proposal & Experimental Results

For uni-directional prediction, the current 8-tap DCT-IF filter coefficients are used as follows:

Half pixel: [-1, 4,-11, 40, 40,-11, 4, -1]

Quarter pixel: [-1, 4,-10, 57, 19, -7, 3, -1]

For bi-directional prediction, the proposal utilizes the 6-tap DCT-IF filter coefficients in CE3 of the Daegu meeting [1] (coefficients are rounded to same 6-bit accuracy of current 8-tap DCT-IF). The coefficients are given as:

Half pixel: [1, -8, 39, 39,-8, 1]

Quarter pixel: [2, -8, 57, 18,-6, 1]

The experimental results are given as follows (It should be noted that the results are obtained in a heterogeneous cluster so the encoding/decoding times are not reliable). The gain in coding efficiency is attributed to slight variation in filter responses for uni- and bi-prediction as seen in the below Figure (i.e. the proposal is gaining partly because encoder now has the capability to adapt the filter).



Figure 1 Frequency response of 8-tap and 6-tap dct-if filter

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Random access LoCo |  |
| Y BD-rate | U BD-rate | V BD-rate |
| Class A | -0.6 | 0.1 | 0.1 |
| Class B | -0.4 | 0.0 | 0.0 |
| Class C | 1.4 | 0.6 | 0.6 |
| Class D | 3.5 | 1.5 | 1.5 |
| Class E |  |  |  |
| All | 0.9 | 0.5 | 0.5 |
| Enc Time[%] | 100% | | |
| Dec Time[%] | 86% | | |
|  |  |  |  |
|  | Low delay LoCo | | |
|  | Y BD-rate | U BD-rate | V BD-rate |
| Class A |  |  |  |
| Class B | -0.6 | 0.2 | 0.4 |
| Class C | 0.8 | 0.4 | 0.3 |
| Class D | 2.0 | 0.5 | 0.7 |
| Class E | -2.0 | -0.6 | -0.1 |
| All | 0.1 | 0.2 | 0.3 |
| Enc Time[%] | 81% | | |
| Dec Time[%] | 72% | | |

One thing to note is that majority of the loss is due to the low resolution BQSquare sequence, and when that sequence is excluded the loss becomes 0. For high-resolution Class-A,B, E sequences, the proposal improves the coding efficiency of low complexity cases by 0.8% on average.

# Complexity

Using the previous CE3 complexity estimate, the complexity of 6-tap DCT-IF filter is found to be around 70% of 8-tap DCT-IF. This means that the worst-case interpolation complexity is reduced by around 30%.

More importantly, the worst case memory bandwidth is also reduced by around 33%, due to the reduced number of filter taps. This reduction is found as follows: The worst-case for memory bandwidth is when the motion vector points to one of the sub-pixels not aligned with horizontal or vertical pixels and motion block size is 4x4. For 8-tap case, decoder needs to load two blocks of 11x11 pixels, but for 6-tap case decoder needs to load 2 blocks of 9x9 reducing the worst case memory bandwidth requirement by around 33%.

# Conclusions

It is proposed to use interpolation filter with fewer taps for bi-prediction. It is proposed to adopt this technique in HM due to the following benefits:

* Both the worst-case interpolation complexity and worst-case memory bandwidth is reduced by around 30%.
* This complexity benefit is coming with a small penalty on coding efficiency. In fact majority of the loss is due to the low resolution BQSquare sequence, and when that sequence is excluded, the loss becomes 0. In addition, for high resolution sequences where HEVC is arguably more relevant, there is a 0.8% coding efficiency improvement on average.
* The proposal is utilizing well-known 6-tap separable filter using the DCT-IF coefficients, which is earlier shown to be implementation friendly.

# References

[1] E. Alshina, J. Chen, A. Alshin, N. Shlyakhov , W.J. Han, “CE3: Experimental results of DCTIF by Samsung”, Joint Collaborative Team on Video Coding, JCTVC-D344r1 Daegu, Jan. 2011¨

# Patent rights declaration(s)

**Nokia may have IPR 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).**