|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  11th Meeting: Shanghai, CN, 10–19 Oct. 2012 | Document: JCTVC-K0325 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Picture Adaptive Frame/Field Coding for HEVC** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Information | | |
| *Author(s) or Contact(s):* | Raul I Lopez RGB Networks, Inc. 390 West Java Drive Sunnyvale, CA 94089 U.S.A.  Arianne T. Hinds  Cable Television Laboratories  858 Coal Creek Circle  Louisville, CO 80027  U.S.A. | Tel: Email:  Tel:  Email: | +1 408 701 2898 [rlopez@rgbnetworks.com](mailto:rlopez@rgbnetworks.com)  +1 303 661 3419  a.hinds@cablelabs.com |
| *Source:* | RGB Networks, Inc., Cable Television Laboratories | | |

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

# Abstract

The document shows results from using HM 8.0 for two progressive and two interlaced sequences encoding each as progressive and interlaced sequence, in addition results from x264 MBAFF Encoding are included too for reference.

Finally, a recommendation for how to modify the syntax and HM code is made to implement Picture-Adaptive Frame Field (PAFF) coding in HEVC.

# Introduction

To show how much better interlaced coding (coding each field of a frame as an independent picture) is versus progressive coding (coding both fields interleaved as a frame) two experiments were run.

# Experimental Setup

## HEVC Frame Encoder

The HM 8.0 release was used encoding each full 1920x1080 frame regardless of its origin from interlaced or progressive video. ME Search Range of 128 was selected, otherwise all other default parameters are used.

## HEVC Field Encoder

The HM 8.0 release was used encoding each half 1920x540 field as a frame regardless of its origin from interlaced or progressive video. ME Search Range of 128 was selected, otherwise all other default parameters are used.

## X264 MBAFF

X264 0.123.x was used to generate MBAFF H.264 Encoded results with this command line:

x264

--input-res 1920x1080

--aud

--output OUTPUT.h264

--qp 41

--fps 29.97

--keyint 32

--tff

--open-gop

--ref 4

--merange 128

--preset slower

--dump-yuv RECON.yuv

--frames 60

INPUT.yuv

The following set of QP values is used: “38, 39, 40, 41.” Thus QPx264 = QPHM+2.

## Test Video Clips

An external source of interlace video was used. NBC Universal (NBCU) provided us with a set of 1920x1080i@60Hz video clips for testing under a content license agreement. The NBCU test set consists of the following sequences. The sequences were converted from 4:2:2 chroma format to 4:2:0 using TM5-based chroma-conversion filters.

Table : NBCU Interlaced Test Video Sequences

| **Name** | **Frames** | **Dur.(s)** | **Content** | **Additional Comments** |
| --- | --- | --- | --- | --- |
| MacysParade\_Flags | 1984 | 66.1 | Macy’s Thanksgiving Day Parade live coverage of a marching band with flag wavers | interlaced, medium camera motion, uni-directional action, saturated colors, high detail, frames 0-59. |
| MacysParade\_Musical | 1421 | 47.4 | Macy’s Thanksgiving Day Parade live coverage of singing and dancing performers including acrobatics | Interlaced including some progressive frames, horizontal fast camera motion, multi-directional action, saturated colors, high detail per field, frames 17-80 |

Due to encoding HW limitations only the first 60/64 frames of the sequences were compared. Because all the test sequences are 30fps, the resulting GOP size is 32 for frame coding and 64 for field coding.

In addition to compare using Field Coding on a Progressive Sequence, the first 60/64 frames of Crowd Run (1080p50) )/Red Kayak (1080p29.97, frames 64-127) were used as well.

# Experimental Results

## Bitstream-Level BD-Rate

|  |  |  |  |
| --- | --- | --- | --- |
| **HEVC Field Coding vs HEVC Frame Coding** | | | |
| **BD-rate (piecewise cubic)** | | | |
| **Clip** | **Y** | **U** | **V** |
| Flags | -57.6% | 0.0% | -55.6% |
| crowd\_run | 14.6% | 0.0% | 0.0% |
| Musical | -66.7% | -65.0% | 0.0% |
| red\_kayak | -1.8% | 22.4% | 25.8% |
| Average | -27.9% | -10.7% | -7.5% |

|  |  |  |  |
| --- | --- | --- | --- |
| **HEVC Field Coding vs x264 MBAFF** | | | |
| **BD-rate (piecewise cubic)** | | | |
| **Clip** | **Y** | **U** | **V** |
| Flags | -58.2% | -51.1% | -43.5% |
| crowd\_run | -30.4% | 0.0% | 0.0% |
| Musical | -59.7% | -51.9% | -49.2% |
| red\_kayak | -29.7% | 0.0% | -37.0% |
| Average | -44.5% | -25.8% | -32.4% |

|  |  |  |  |
| --- | --- | --- | --- |
| **HEVC Frame Coding vs x264 MBAFF** | | | |
| **BD-rate (piecewise cubic)** | | | |
| **Clip** | **Y** | **U** | **V** |
| Flags | -4.8% | 0.0% | 0.0% |
| crowd\_run | -37.5% | -39.0% | -38.2% |
| Musical | 20.8% | 0.0% | 0.0% |
| red\_kayak | -28.7% | -27.5% | -47.7% |
| Average | -12.5% | -16.6% | -21.5% |

## Analysis of Results

The results above show that for “Flags”, the interlaced sequence, HEVC Field Coding performed best against both HEVC Frame Coding and x264. Comparing HEVC Frame Coding vs x264 MBAFF showed almost identical performance.

For “crowd\_run”, the progressive sequence, HEVC Field Coding performed worse than HEVC Frame Coding but better than x264 MBAFF.

For “Musical”, HEVC Field Coding performed best, followed by x264 MBAFF and HEVC Frame Coding last.

Finally, for “red\_kayak”, HEVC Field Coding was marginally better than HEVC Frame Coding in Luma but not in Chroma and both better than x264 MBAFF.

The overall averages for each experiment point to HEVC Field coding as most effective for video of interlaced origin or with high uncorrelated motion (Red Kayak running water). The problem with choosing HEVC Field or Frame coding arises when Transcoding video material that originates from MPEG-2 or H.264 with macroblok-level Field/Frame decision.

# Picture Adaptive Frame/Field (PAFF) Coding for HEVC

## Complexity

To be able to decide on encoding a frame or two fields, the target frame/fields need to be compared versus the relevant reference(s) presented as both frame and fields. To be thorough, 3X the number of comparisons need to be done: fields of same polarity, fields of opposite polarity and full frame. An optimization is to only compare fields of the same parity for a 2X increase in comparisons. A comparison can be a SAD done at the maximum CU width/height, the sum of SADs with the smaller number would be the preferred way to go: Frame or Fields (same or opposite polarity).

## Syntax Changes

To accommodate the ability of each frame to be progressive or interlaced the following high-level changes are needed:

### Remove “field\_seq\_flag” from vui\_parameters.

### Add “pict\_struct\_present\_flag” to vui\_parameters.

### Flag using SEI messages, per frame, whether a frame should be treated as a progressive frame or fields.

## HM Estimate for implementing PAFF for HEVC.

### Close on HEVC Syntax Changes (2 weeks)

### Examine HM for places to change (2 weeks)

### Implement HM changes (3 weeks)

### Verify Changes (6 weeks)

# Summary and Recommendations

More sequences are needed to draw more relevant conclusions but adding Hoang and Xin’s recommendations and analysis, we recommend to study the addition of the PAFF mode to HEVC.

# References

1. D. Hoang and J. Xin, “Coding interlaced video using sequence-adaptive field-frame,” JCT-VC Document JCTVC-K0250, Shanghai, CN, Oct. 2012.

# Patent rights declaration(s)

**RGB Networks, Inc., does not have any current or pending patent rights relating to the technology described in this contribution.**

**Cable Television Laboratories does not have any current or pending patent rights relating to the technology described in this contribution.**