|  |  |
| --- | --- |
| **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-K0193\_r1 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **AHG7: Adaptive colour-space transformation of residual signals** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Kei Kawamura Tomonobu Yoshino Sei Naito  2-1-15, Ohara, Fujimino-shi, Saitama, JAPAN | Tel: Email: | +81 49 278 7411 ki-kawamura@kddi.com |
| *Source:* | KDDI Corp. (KDDI R&D Laboratories, Inc.) | | |

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

# Abstract

This contribution proposes the adaptive color-space transformation of residual signals for HEVC FrExt. This approach adaptively transforms prediction error signals from 4:4:4 chroma format to an sub-optimal colour space which reduce redundancy of pixel value between planes. The transformation matrix is derived for each coding unit by a singular-value-decomposition. The colour-space transformation is applied to prediction error of both intra and inter mode. Simplified singular-value-decomposition is realized only by integer multiplication and shift operation with limited iterations. Additional information is not necessary because residual signals are always transformed. The BD-rate gains for YUV 444 sequences are 6.7%/7.2%/8.7% for all intra / random access / low delay B with 10bit configuration, respectively.

# Introduction

This contribution proposes new color-space transformation steps to HEVC of 4:4:4 chroma format. Transformed target is intra/inter residual signals. Transform matrices are derived from neighboring pixels / reference signals by using singular value decomposition with finite computation steps.

# Proposed method

## Overview

Block diagram of the proposed adaptive- color-space transformation is illustrated in Fig. 1. Compared to the existing HEVC block diagram, matrix derivation and color space transformation/inverse transformation are added. Matrix derivation has one input from either local decoded image or frame buffer and has one output for transform matrix of color space. After intra/inter prediction, residual signals are transformed by the derived matrix. A processing flow of resulting signals is the same as HEVC scheme. In the following section, we describe the each component.



## Transform and inverse transform

In the forward transform, three planes are processed as follows,

In the inverse transform, three components are processed as follows,

Transformed values are clipped within the range of the HEVC specification because the values are enlarged of square of three times in the worst case.

## Derivation algorithm of transformation matrix

Transform matrix is derived from the reference pixel values. Reference pixels are different between intra cases and inter cases.

For intra block, target block and reference pixels are shown in Fig. 1. In this figure, target block consists of 8x8 gray pixels and reference pixel is orange and purple pixels.



Actual reference pixels depend on the intra prediction mode. Orange pixels are referred to when the range of intra prediction mode is from two to ten, while purple pixels are referred to when the range of intra prediction mode is from 26 to 34. Otherwise, both orange and purple pixels are referred to. In any case, the number of reference pixels is the power of two. Therefore upper left region is not included due to the violation of shift operation to compute an average.

For inter block, reference pixel of matrix derivation is the same as that of motion compensation. In order to realize the shift operation, some kind of AMP block is sub-sampled. For example, number of reference pixels in a 12x16 block is reduced to 2/3.

After getting the reference pixels, an average of pixels in the same plane is subtracted from the reference pixel values. A covariance matrix of three planes is computed from the subtracted values. For example, when the number of reference pixels is 32, a covariance matrix is computed by . The covariance matrix is then normalized by the shit operation so that maximum value of diagonal elements holds within the range from 4096 to 8191.

After generating normalized covariance matrix, singular de-convolution is applied based on Jacob algorithm. A maximum number of main iteration is decided to three. A square root operation of arbitrary number is also realized by integer operation with two iterations. Finally, each column is sorted so that diagonal elements value is decreased and then transposed. The obtained matrix is named .

When the reference pixels are not available, utilizes a unit matrix.

## Modification of coding structure

Coding structure is required to be identical among all planes because of the color space transformation. For instance, following items need the same properties; the size of coding unit, the type of prediction unit, the mode of intra prediction and the depth of transform unit.

An influence of these modifications is studied in JCTVC-K0192 because this modification is generally useful for 444 chroma format.

## Rate-distortion optimization

Rate distortion optimization (RDO) for intra frame is modified in order to decide the coding unit and prediction mode for all planes simultaneously. For inter frame, all plane have been equally considered. This RDO modification has no influence to the RDOQ process.

## Syntax

Any syntax modification in not required by this scheme. Coding structure modification and tools/mode limitations are however occurred. This modification is based on JCTVC-K0xx3.

# Experimental results

The proposed mode is implemented to HM8.0/Ahg7.

Common test condition for AHG7 is follows as the BoG report of JCTVC-J0581.

Table x shows the summary of BD-rate for YUV444 sequences.

Table 1 Results of YUV444 sequences

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra HE10** | | | **Random Access HE10** | | | **Low delay B HE10** | | |
|  | Y | U | V | Y | U | V | Y | U | V |
| Kimono | -6.2% | 0.4% | -1.2% | -6.5% | -1.7% | -0.6% | -6.8% | -0.7% | -1.0% |
| Parkscene | -8.1% | -6.9% | -2.7% | -8.1% | -8.3% | -6.0% | -9.6% | -7.0% | -6.5% |
| BirdsInCage | -7.9% | 8.4% | -11.9% | -8.4% | 27.4% | -13.7% | -16.9% | 15.2% | -32.3% |
| DucksAndLegs | -9.4% | 5.5% | -0.2% | -15.9% | 10.6% | -12.8% | -13.8% | 11.8% | -12.3% |
| Traffic | -2.1% | 1.6% | -1.7% | -4.2% | 9.2% | -0.5% | -6.0% | 9.5% | -1.8% |
| CrowdRun | -4.5% | -7.4% | -7.7% | -3.4% | -8.8% | -6.8% | -3.1% | -8.0% | -6.8% |
| OldTownCross | -3.8% | -1.0% | -2.1% | -2.5% | 4.4% | 2.3% | -5.5% | 5.2% | -3.6% |
| Seeking | -11.5% | -17.0% | -12.6% | -8.6% | -21.1% | -9.2% | -8.2% | -19.2% | -9.3% |
| **Overall** | -6.7% | -2.1% | -5.0% | -7.2% | 1.5% | -5.9% | -8.7% | 0.9% | -9.2% |
| Enc Time[%] | 147% | | | 111% | | | 109% | | |
| Dec Time[%] | 106% | | | 115% | | | 116% | | |

Following table shows the summary of BD-rate for additional RGB444 sequences, which are obtained from VQEG web site. The proposed method is effective for RGB444 sequences.

Table 2 Results of RGB444 sequences

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra HE10** | | | **Random Access HE10** | | | **Low delay B HE10** | | |
|  | G | B | R | G | B | R | G | B | R |
| CrowdRun | -31.5% | -21.3% | -24.5% | -42.9% | -20.5% | -26.3% | -42.7% | -15.8% | -22.9% |
| ParkJoy | -19.6% | -8.7% | -14.7% | -33.1% | -1.7% | -25.4% | -29.4% | -1.5% | -21.0% |
| DucksTakeOff | -22.6% | -15.3% | -13.9% | -36.0% | 0.0% | -13.0% | -36.9% | -6.4% | -10.2% |
| InToTree | -22.4% | -8.9% | -9.9% | -40.2% | 6.7% | -2.8% | -47.0% | 9.0% | -10.7% |
| OldTownCross | -25.4% | -17.7% | -22.3% | -32.3% | -15.8% | -26.9% | -31.3% | -13.8% | -24.7% |
| **Overall** | -24.3% | -14.4% | -17.1% | -36.9% | -6.3% | -18.9% | -37.5% | -5.7% | -17.9% |
| Enc Time[%] | 139% | | | 104% | | | 104% | | |
| Dec Time[%] | 98% | | | 107% | | | 105% | | |

Following tables show the summary of BD-rate with chroma clipping for YUV444 and RGB444 sequences. Influence of the chroma clipping is described in JCTVC-K0192.

Table 3 Results of YUV444 sequences

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra HE10** | | | **Random Access HE10** | | | **Low delay B HE10** | | |
|  | Y | U | V | Y | U | V | Y | U | V |
| Kimono | -1.8% | -1.2% | -6.0% | -2.4% | -7.3% | -8.1% | -3.3% | -6.6% | -9.0% |
| Parkscene | -3.9% | -8.9% | -6.7% | -3.7% | -13.2% | -11.7% | -6.0% | -12.9% | -12.5% |
| BirdsInCage | -4.9% | 7.6% | -15.6% | -1.4% | 19.9% | -19.6% | -11.6% | 11.5% | -39.2% |
| DucksAndLegs | -4.1% | 4.1% | -2.7% | -5.2% | 9.0% | -15.9% | -7.0% | 9.6% | -17.7% |
| Traffic | 1.3% | -1.8% | -6.0% | 1.8% | 3.8% | -6.2% | 0.1% | 4.1% | -8.2% |
| CrowdRun | -1.8% | -10.9% | -11.8% | -0.4% | -16.1% | -14.5% | -0.5% | -15.0% | -14.5% |
| OldTownCross | -1.9% | -1.3% | -4.4% | 1.6% | 2.3% | -5.9% | -3.8% | 3.8% | -12.0% |
| Seeking | -9.0% | -17.9% | -15.4% | -5.4% | -25.1% | -16.3% | -5.6% | -22.7% | -16.3% |
| **Overall** | -3.3% | -3.8% | -8.6% | -1.9% | -3.3% | -12.3% | -4.7% | -3.5% | -16.2% |
| Enc Time[%] | 149% | | | 112% | | | 110% | | |
| Dec Time[%] | 108% | | | 116% | | | 117% | | |

Table 4 Results of RGB444 sequences

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **All Intra HE10** | | | **Random Access HE10** | | | **Low delay B HE10** | | |
|  | G | B | R | G | B | R | G | B | R |
| CrowdRun | -29.1% | -21.4% | -24.6% | -39.5% | -20.4% | -26.3% | -39.6% | -16.1% | -23.2% |
| ParkJoy | -16.6% | -8.6% | -14.3% | -28.9% | -3.1% | -24.2% | -25.4% | -2.6% | -19.9% |
| DucksTakeOff | -19.8% | -13.6% | -14.8% | -28.5% | 0.7% | -11.9% | -27.9% | -5.1% | -10.6% |
| InToTree | -18.9% | -8.4% | -10.5% | -31.2% | 6.9% | -1.4% | -37.7% | 8.7% | -8.3% |
| OldTownCross | -22.6% | -18.3% | -22.1% | -28.4% | -17.5% | -26.5% | -27.7% | -15.7% | -24.4% |
| **Overall** | -21.4% | -14.1% | -17.2% | -31.3% | -6.7% | -18.1% | -31.7% | -6.2% | -17.3% |
| Enc Time[%] | 144% | | | 109% | | | 108% | | |
| Dec Time[%] | 102% | | | 112% | | | 110% | | |

# Conclusion

This contribution proposed adaptive color-space transformation of residual signals. The proposed method reduced the redundancy of color space in 4:4:4 chroma format. Since transformation matrix was derived from sample values on both encoder and decoder side, side information was not necessary. On the other hand, actual transformed target are signals in residual domain. Computation of matrix derivation is completely realized by integer multiplication and shift operation with limited iterations.

Experimental results showed that the proposed method achieveed 6.7%, 7.2%, and 8.7% BD-rate reduction respectively for all intra, random access, and low-delay B configuration with HE10 for YUV444 sequences.

# References

1. D. Flynn, “BoG report: Extended chroma formats,” JCTVC-J0581, Stockholm, July 2012.

# Patent rights declaration(s)

**KDDI Corporation 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).**