|  |  |
| --- | --- |
| **Joint Collaborative Team on 3D Video Coding Extensions**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  11th Meeting: Geneva, CH, 12–18 Feb. 2015 | Document: JCT3V-K0028 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Encoder optimization: Motion estimation for DBBP with masked SAD instead of SATD** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Fabian Jäger Institut für Nachrichtentechnik RWTH Aachen University  Jacek Konieczny, Giovanni Cordara Huawei Technologies Munich Office, European Research Center | Tel: Email: | +49 (0) 241 80 27678 [jaeger@ient.rwth-aachen.de](mailto:jaeger@ient.rwth-aachen.de)  +49 (0) 89 158834 4334 [jacek.konieczny@huawei.com](mailto:jacek.konieczny@huawei.com) |
| *Source:* | RWTH Aachen University, Huawei Technologies | | |

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

# Abstract

At the 6th JCT-3V meeting, a motion/disparity prediction method was introduced in JCT3V-G0106 [1] that uses a depth-derived binary segmentation mask for the derivation of PU partitioning and for merging of two prediction signals. The method is called Depth-based Block Partitioning (DBBP). The segmentation process in JCT3V-G0106 results in two segments, for which two sets of motion parameters need to be estimated at the encoder side. Motion estimation for DBBP is performed for each segment by evaluating a masked distance measure (e.g. SATD). As SATD computes the error signal in the transform domain, this masking operation does not result in the intended optimization criterion. Consequently, this document proposes to always use a masked SAD criterion when performing motion estimation for DBBP. This encoder optimization yields 0.3% coding efficiency improvements compared to the latest anchor HTM-13.0 without any impact on the encoder complexity.

# Segment-wise Motion Estimation

By performing two full block compensations with independent motion parameters, motion or disparity discontinuities can be modeled. At the encoder side the two sets of motion parameters need to be estimated for a block utilizing DBBP. An encoder’s motion estimation algorithms try to find a rectangular region in a reference picture of the same size as the current prediction block, which minimizes a distortion metric, such as the SATD. With arbitrarily shaped segments this approach needs to be adapted to result in decent estimations of optimal motion parameters for each segment. The basic approach of extending block-based matching algorithms to arbitrarily shaped segments is to invalidate all sample positions within the current block, which do not belong to the currently active segment. *Invalidation* in this context means that these sample positions are ignored when computing the distortion in the block matching algorithm’s optimization loop. Consequently, the distortion for a particular candidate motion parameter set is computed based on the currently active segment’s samples only. The rationale behind this method is that potential distortions at the other segment’s sample positions do not impact the overall block’s distortion as these areas are overwritten during the subsequent merging step of DBBP.

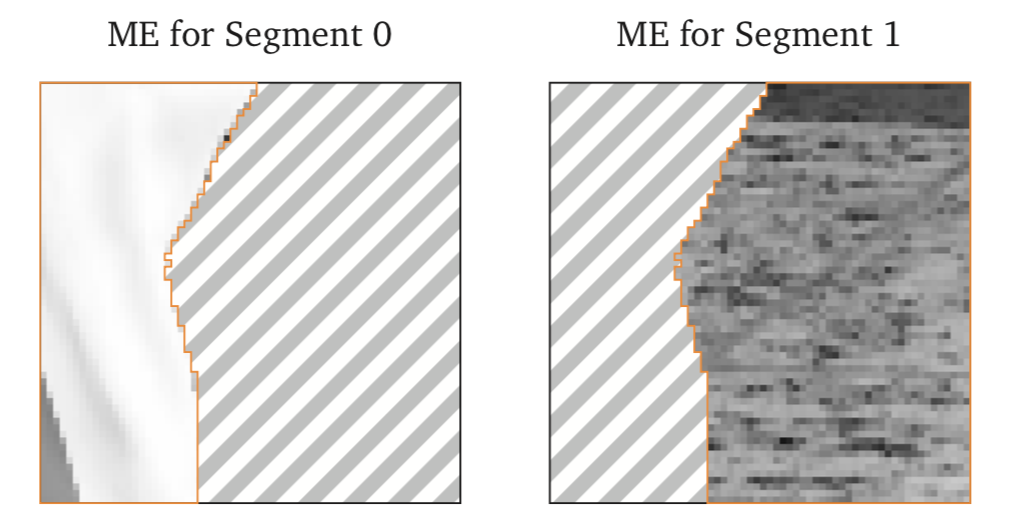


Figure : Illustration of the modified motion estimation algorithm for DBBP. For the active segment (orange), only the corresponding sample positions are included in the computation of the distortion. The gray, hatched region of the other segment is ignored when computing the SAD.

This minor modification of the encoder does not require further changes to the motion estimation process and allows finding optimal motion parameters for each arbitrarily shaped segment by reusing the block-based matching algorithms. Modern encoders utilize SATD to measure the distortion for each candidate reference block. SATD transforms the block of residual values by the Hadamard transform before the SAD is computed in the transform domain. Due to this additional transform step the aforementioned invalidation of sample positions is not combinable with the SATD, because invalidated positions are not transferrable to the transform domain. The most straight-forward solution to this incompatibility is to disable SATD for motion estimation for DBBP blocks. Whenever the encoder tries to find the optimal motion parameters for an DBBP block, the SAD (in the spatial domain) is used for distortion measurement. Thus, the described invalidation of sample positions still works for the estimation of motion parameters for arbitrarily shaped segments. For all other (existing) motion or disparity prediction modes the encoder still uses SATD for its block-wise motion parameter search.

For the final decision on the optimal coding mode for a particular CU, the resulting rate-distortion cost for the DBBP mode is still compared against the final rate-distortion costs of the conventional modes. At this optimization stage, the corresponding distortion term is always computed based on the SAD and therefore the encoder does not need to treat DBBP differently to decide on the optimal coding configuration for a CU.

# Simulation Results

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | video 0 | video 1 | video 2 | video PSNR / video bitrate | video PSNR / total bitrate | synth PSNR / total bitrate | enc time | dec time |
| Balloons | 0.0% | -0.4% | -0.3% | -0.2% | -0.1% | -0.1% | 101.0% | 97.1% |
| Kendo | 0.0% | -0.5% | -0.4% | -0.2% | -0.2% | -0.1% | 97.3% | 107.9% |
| Newspaper\_CC | 0.0% | -0.5% | -0.4% | -0.2% | -0.1% | 0.0% | 101.2% | 99.3% |
| GT\_Fly | 0.0% | -0.8% | -0.8% | -0.2% | -0.2% | -0.1% | 100.3% | 103.7% |
| Poznan\_Hall2 | 0.0% | -0.3% | -0.1% | -0.1% | -0.1% | -0.2% | 105.4% | 101.8% |
| Poznan\_Street | 0.0% | -0.6% | -0.4% | -0.2% | -0.1% | -0.1% | 100.4% | 100.3% |
| Undo\_Dancer | 0.0% | -3.1% | -3.3% | -0.9% | -0.9% | -0.6% | 101.9% | 109.9% |
| Shark | 0.0% | -1.3% | -1.0% | -0.3% | -0.3% | -0.2% | 100.3% | 100.1% |
| 1024x768 | 0.0% | -0.4% | -0.4% | -0.2% | -0.2% | -0.1% | 99.8% | 101.4% |
| 1920x1088 | 0.0% | -1.2% | -1.1% | -0.4% | -0.3% | -0.2% | 101.7% | 103.2% |
| **average** | **0.0%** | **-0.9%** | **-0.8%** | **-0.3%** | **-0.3%** | **-0.2%** | **101.0%** | **102.5%** |

# Cross Check

The cross check was performed by NTT. They investigated the proposed changes to the software and the specification text and ran the same simulation configuration to confirm the presented simulation results.

# Conclusion

In this contribution it is proposed to always use a masked SAD operation for the distance or distortion metric when performing motion estimation for DBBP. This encoder-only modification yields about 0.3% coding efficiency improvements over the latest anchor HTM-13.0 without any changes to the encoder complexity. Consequently, this proposal seems to be a reasonable modification to the encoder software package.

# Patent rights declaration(s)

**RWTH Aachen University 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).**

**Huawei Technologies 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).**