**Joint Video Experts Team (JVET)**
**of ITU-T SG21 WP3/21 and ISO/IEC JTC 1/SC 29**
38th Meeting, by teleconference, 26 March – 4 April 2025

Document: JVET-AL0199-v2

| | |
|---|---|
| *Title:* | **EE2-6.1: Adaptive picture-level vertical mirroring** |
| *Status:* | Input document to JVET |
| *Purpose:* | Proposal |

*Author(s) or Contact(s):* Dawid Mieloch, Jakub Stankowski, Mateusz Lorkiewicz, Adrian Dziembowski, Damian Karwowski

Polanka 3, 61-131 Poznań, Poland

Tel:
Email: dawid.mieloch@put.poznan.pl

*Source:* Poznan University of Technology (PUT)
*Institute of Multimedia Telecommunications*


# Abstract

The document reports the results of EE2-6.1 in improving video coding by introducing adaptive picture-level vertical mirroring. The method involves an automatic frame pre-processing step for determining whether the mirroring should be applied without the necessity to perform a brute force search or apply additional RDO checks. In EE2-6.1 the method was additionally updated to work in RA configuration. The proposed method was implemented in the ECM-16.1 software and achieved an average BD-rate change of -0.22%, -0.25%, and -0.43% for Y, U, and V components in AI configuration, and -0.11%, -0.12%, and -0.15% for RA.


# 1 Introduction

As our other document shows [JVET-AK0171], it can be beneficial, in terms of coding efficiency, to rotate or mirror the input frame before encoding. As was observed in the results presented in the mentioned document, most of the gain comes from vertical mirroring, while other transformations provide much more minor improvements. Additionally, besides vertical mirroring and the 180 rotation, in other transformations, the phase of chroma does not match their collocated luma, requiring additional corrections to be performed. Therefore, the proposal described below provides the automatic method for picture-level vertical mirroring.

# 2 Proposed method

## *2.1 Details of the previous meeting proposal (JVET-AK173)*

The selection of optimal transform is performed before encoding the whole picture. No additional in-loop optimizations have to be performed, and no additional coding modes are introduced in the computational graph, reducing the computational burden of an encoder.

Due to using only the vertical mirroring, the phase of chroma samples still matches their collocated luma samples, so no change is required.

Input image $I_{org}$ (each component independently) is filtered using a 3x3 Sobel filter for 45-degree edges detection (resulting in image $I_{45}$) and a 3x3 Sobel filter for 135-degree edges detection (resulting in image $I_{135}$).

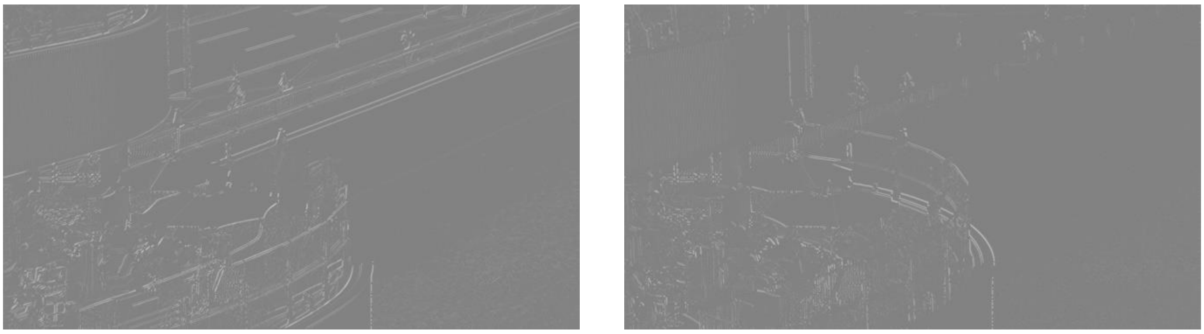

Fig. 1. The result of the 3x3 Sobel filter for 45-degree (left) and 135-degree (right) edges detection.

For each pixel (x,y), it is checked whether one of the edge directions is dominant, i.e., the difference between $I_{45}(x,y)$ and $I_{135}(x,y)$ greater than a threshold (set, e.g., to 160 for a 10-bps video). If so, only the dominant pixel value persists, and the other is set to zero.

Then, all the non-zero values of $I_{45}$ and $I_{135}$ are counted:

$$S_{45} = \sum_{x,y} \left| I_{45}^{Y}(x,y) \right| > 0 \mathbin{?} 1 : 0 + \sum_{x,y} \left| I_{45}^{Cb}(x,y) \right| > 0 \mathbin{?} 1 : 0 + \sum_{x,y} \left| I_{45}^{Cr}(x,y) \right| > 0 \mathbin{?} 1 : 0\,,$$

$$S_{135} = \sum_{x,y} \left| I_{135}^{Y}(x,y) \right| > 0 \mathbin{?} 1 : 0 + \sum_{x,y} \left| I_{135}^{Cb}(x,y) \right| > 0 \mathbin{?} 1 : 0 + \sum_{x,y} \left| I_{135}^{Cr}(x,y) \right| > 0 \mathbin{?} 1 : 0\,.$$

If $S_{135} > (S_{45} + TH)$, the frame is not mirrored. Otherwise, it is mirrored vertically. The threshold is introduced to give the possibility of favoring the selection of not changing the image. In the presented results, TH was set to -6000.

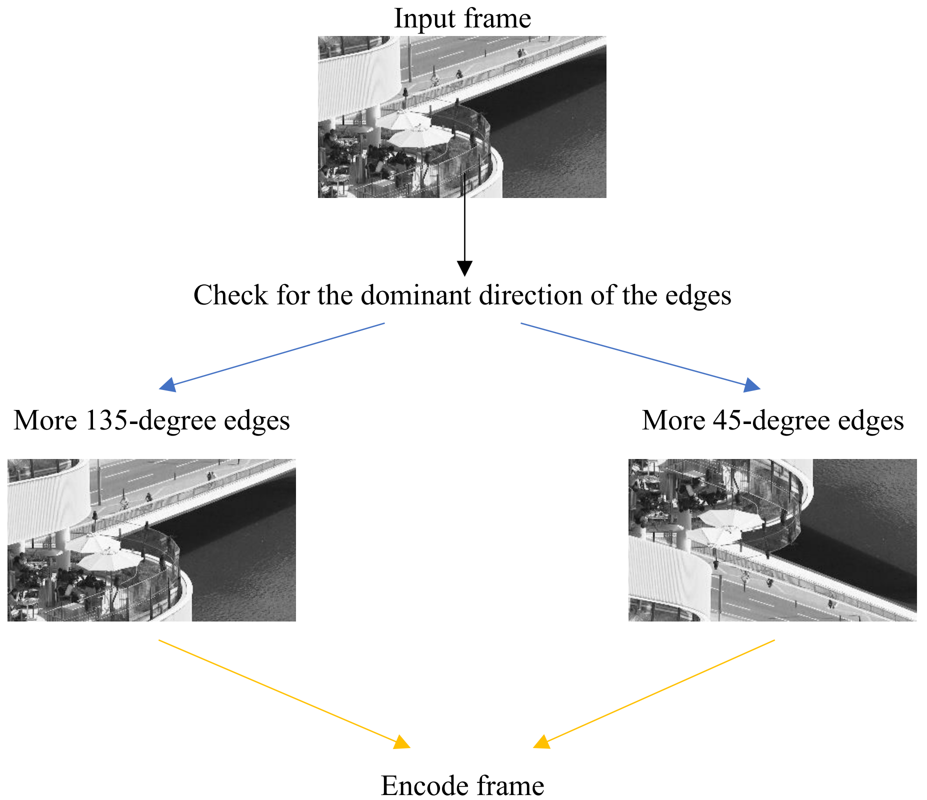


Fig. 2. The overall scheme of the encoder in the proposal.

Naturally, after decoding the frame, the original orientation of the image is restored. The signaling for each frame, in this case, is limited to only 1 bit for indicating the use of the method and 1 bit for signaling that the frame was mirrored.

## *2.2 Details of EE2-6.1 implementation*

The implementation operates exclusively on intra-coded frames and involves a frame analysis phase characterized by simplified filters. This phase does not necessitate memory allocation and maintains low computational complexity. Following the analysis, a frame transform phase is executed, which also exhibits low complexity and performs an in-place picture transform. Similar to the analysis phase, this transform phase does not require memory allocation and operates solely on existing buffers. The schemes of the encoder and the decoder are presented in Fig. 3 and Fig. 4.

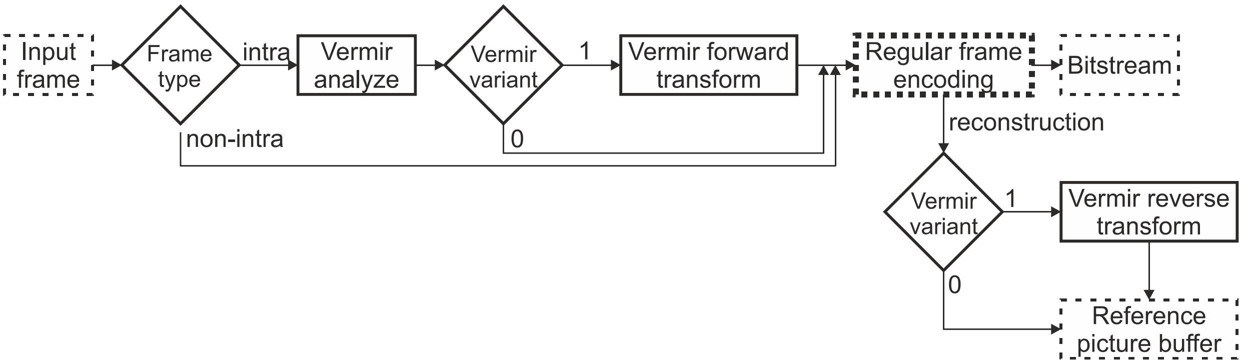


Fig. 3. Encoder block diagram (vermir – vertical mirroring).

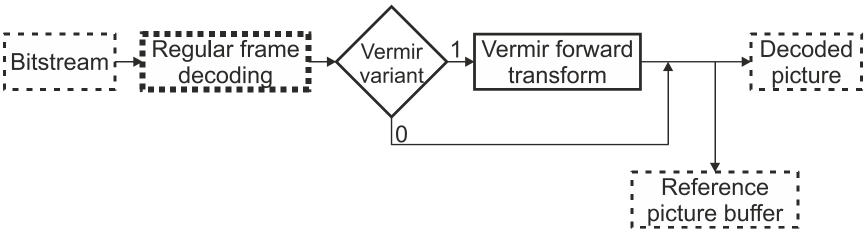


Fig. 4. Decoder block diagram (vermir – vertical mirroring).

# 3 Experimental results

In Table 1, the CTC-based results for AI configuration are presented. Table 2 presents the results for RA. In v4, the decoder for AI was run again on one machine, making the measurement more reliable.

Table 1. CTC-based results for AI configuration.

| | All Intra Main 10 | | | | | | |
|---|---|---|---|---|---|---|---|
| | Over ECM-16.1 | | | | | | |
| | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 | -0.19% | -0.07% | -0.27% | 102.7% | 98.8% | 100.0% | 100.1% |
| Class A2 | -0.21% | 0.10% | -0.07% | 102.5% | 98.4% | 100.0% | 100.0% |
| Class B | -0.58% | -0.92% | -1.32% | 100.1% | 101.2% | 99.1% | 100.3% |
| Class C | 0.01% | -0.01% | -0.02% | 101.4% | 99.2% | 100.4% | 100.0% |
| Class E | 0.01% | 0.01% | 0.01% | 108.1% | 99.1% | 100.0% | 100.0% |
| Overall | -0.22% | -0.25% | -0.43% | 102.5% | 99.6% | 99.9% | 100.1% |
| Class D | 0.01% | 0.01% | 0.01% | 109.4% | 98.9% | 100.1% | 100.0% |
| Class F | 0.09% | 0.41% | 0.62% | 103.3% | 100.0% | 99.0% | 100.0% |

Table 2. CTC-based results for RA configuration.

| | Random Access Main 10 | | | | | | |
|---|---|---|---|---|---|---|---|
| | Over ECM-16.1 | | | | | | |
| | Y | U | V | EncT | DecT | EncVmPeak | DecVmPeak |
| Class A1 | -0.08% | 0.17% | 0.18% | 100.4% | 86.4% | 100.1% | 100.0% |
| Class A2 | 0.01% | 0.12% | 0.06% | 99.5% | 100.4% | 99.7% | 100.8% |
| Class B | -0.29% | -0.54% | -0.64% | 94.2% | 108.3% | 100.0% | 100.0% |
| Class C | 0.00% | 0.00% | 0.04% | 96.1% | 100.5% | 100.0% | 86.7% |
| Class E | | | | | | | |
| Overall | -0.11% | -0.12% | -0.15% | 97.0% | 100.0% | 100.0% | 96.4% |
| Class D | 0.00% | 0.00% | 0.00% | 93.1% | 90.4% | 100.0% | 100.0% |
| Class F | -0.10% | 0.04% | -0.19% | 93.8% | 93.8% | 101.1% | 99.8% |

Table 3. The percentage of frames mirrored vertically (AI configuration).

| | |
|---|---|
| Tango2 | 81% |
| FoodMarket4 | 11% |
| Campfire | 76% |
| CatRobot | 16% |
| DaylightRoad2 | 61% |
| ParkRunning3 | 53% |
| MarketPlace | 53% |
| RitualDance | 57% |
| Cactus | 6% |
| BasketballDrive | 6% |
| BQTerrace | 0% |
| BasketballDrill | 0% |
| BQMall | 0% |
| PartyScene | 0% |
| RaceHorses | 29% |
| BasketballPass | 0% |
| BQSquare | 0% |
| BlowingBubbles | 0% |
| RaceHorsesLow | 0% |
| FourPeople | 0% |
| Johnny | 0% |
| KristenAndSara | 0% |
| BasketballDrillText | 0% |
| ArenaOfValor | 75% |
| SlideEditing | 58% |
| SlideShow | 19% |

# 4 Conclusions

This exploration experiment has explored the potential of picture-level vertical mirroring as a pre-processing step to enhance video coding efficiency. The results demonstrate notable improvements in coding efficiency, especially for A1, A2, and B classes.

# 5 Recommendations

It is recommended to include the possibility of using adaptive picture-level frame mirroring in ECM 17

# 6 Patent rights declaration(s)

**Poznan University of Technology 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).**