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-AL0202

Title:	**EE2-related: Adaptive picture-level vertical mirroring based on decimated video**
Status:	Input document to JVET
Purpose:	Proposal

Author(s) or Contact(s):	Dawid Mieloch, Mateusz Lorkiewicz, Adrian Dziembowski 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 proposes a method for improving video coding by introducing adaptive picture-level vertical mirroring based on encoding of decimated video. The method involves a frame pre-encoding step to determine whether the mirroring should be applied. The proposed method was tested using the ECM-15.0 software and achieved an average BD-rate change of -0.22%, -0.30%, and -0.41% for Y, U, and V components in AI configuration, with an estimated <2% increase of encoding runtime and no change in a decoder runtime.

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

The selection of optimal transform is performed before encoding the whole picture and is based on brute-force encoding of highly decimated video.

The proposal is an *a posteriori* selection of the best mode for vertical mirroring. Therefore, before main encoding, all frames were encoded two times – without transformation and with vertical mirroring, but using video decimated for this step. Then, the selection was made for video without decimation, for each frame independently to optimize the Y BD-rate.

For example, when videos were decimated 4 times in each direction, we can assume 16 times faster encoding, what gives around 12.5% in overall encoding time increase

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.

3 Experimental results

In Table 1, the CTC-based results for AI configuration are presented. In this experiment the videos were not decimated, or decimated 2/4 times in each direction. The encoding time does not reflect measured runtime but is estimated from the size of the encoded frame.

Table 1. CTC-based results of the proposal in comparison with ECM 15.0 (encoding time estimated).

No decimation

	All Intra Main 10		
	Over ECM-15.0		
	Y	U	V
Class A1	-0.23%	0.00%	-0.18%
Class A2	-0.82%	-0.02%	-0.27%
Class B	-0.89%	-0.70%	-0.89%
Class C	-0.03%	0.06%	0.04%
Class E	0.00%	0.00%	0.01%
Overall	-0.43%	-0.18%	-0.31%

100% encoding time increase

Decimated 2 times

	All Intra Main 10		
	Over ECM-15.0		
	Y	U	V
Class A1	0.15%	0.23%	0.12%
Class A2	-0.67%	-0.03%	-0.18%
Class B	-0.60%	-1.04%	-1.19%
Class C	0.11%	0.12%	0.05%
Class E	0.08%	0.11%	0.10%
Overall	-0.22%	-0.21%	-0.31%

50% encoding time increase

Decimated 4 times

	All Intra Main 10		
	Over ECM-15.0		
	Y	U	V
Class A1	0.00%	0.12%	0.02%
Class A2	-0.49%	-0.20%	-0.29%
Class B	-0.54%	-1.04%	-1.32%
Class C	0.21%	0.19%	0.12%
Class E	0.35%	0.27%	0.35%
Overall	-0.13%	-0.22%	-0.33%

12.5% encoding time increase

Table 2 presents the results of the experiment in which the videos used to determine the optimal transform were scaled to a fixed size (64×64, 128×128, or 256×256). These results show that significant gains can be achieved with less than 2% increase in encoding time. Moreover, the encoding time overhead becomes smaller when the resolution of the original video becomes larger.

64×64

	All Intra Main 10			Estimated encoder runtime
	Over ECM-15.0			
	Y	U	V	
Class A1	-0.04%	0.02%	-0.02%	100.05%
Class A2	-0.18%	0.16%	-0.08%	100.05%
Class B	-0.47%	-0.87%	-1.16%	100.20%
Class C	0.18%	0.16%	0.07%	101.03%
Class E	0.39%	0.26%	0.42%	100.44%
Overall	-0.06%	-0.13%	-0.25%	100.11%

128×128

	All Intra Main 10			Estimated encoder runtime
	Over ECM-15.0			
	Y	U	V	
Class A1	-0.13%	-0.03%	-0.16%	100.20%
Class A2	-0.26%	0.10%	-0.17%	100.20%
Class B	-0.56%	-1.01%	-1.35%	100.79%
Class C	0.13%	0.07%	0.08%	104.10%
Class E	0.13%	0.14%	0.29%	101.78%
Overall	-0.17%	-0.23%	-0.36%	100.46%

256×256

	All Intra Main 10			Estimated encoder runtime
	Over ECM-15.0			
	Y	U	V	
Class A1	-0.16%	-0.08%	-0.09%	100.79%
Class A2	-0.42%	-0.12%	-0.29%	100.79%
Class B	-0.57%	-1.07%	-1.33%	103.16%
Class C	0.12%	0.12%	0.05%	116.41%
Class E	0.04%	0.03%	0.05%	107.11%
Overall	-0.22%	-0.30%	-0.41%	101.83%

4 Conclusions

This document has explored the potential of picture-level vertical mirroring as a pre-processing step that uses encoding of highly decimated video to enhance video coding efficiency. The results demonstrate notable improvements in coding efficiency.

5 Recommendations

This document shows supplementary data in order to support the recommendation from document [AL0199] on EE2-6.1. Therefore, it is recommended to include the possibility of adaptive picture-level frame mirroring in ECM 17.0.

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).