Joint Collaborative Team on Video Coding (JCT-VC)
of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11
4th Meeting: Daegu, KR, 20-28 January, 2011

Document: JCTVC-D213r1

Title:	**In-loop reference frame denoising in HEVC reference software**
Status:	Input Document to JCT-VC
Purpose:	Proposal

Author(s) or Contact(s):	Peter Amon, Andreas Hutter Siemens Corporate Technology Munich, Germany	Tel: Email:	+49 89 636 54642 p.amon @ siemens.com
	Eugen Wige, André Kaup University of Erlangen-Nuremberg Multimedia Communications and Signal Processing Erlangen, Germany	Tel: Email:	+49 9131 85 27104 wige @ lnt.de
Source:	Siemens, University of Erlangen-Nuremberg		

Abstract

This proposal presents an algorithm for in-loop denoising of the reference frame. The algorithm modifies the temporal predictor while the decoded picture is unchanged. Knowledge of the noise power within the reference frame is used in order to improve the inter frame prediction. For noise filtering of the reference frame, a low-complexity denoising algorithm is implemented. In JCT-VC219, the results for the implementation of the algorithm in the H.264/AVC reference software (JM 15.1) were presented. It was shown that the bitrate can be decreased for (high resolution) noisy image sequences especially for higher qualities at medium to high data-rates. This contribution shows that the gains are nearly preserved when implementing the scheme in the HEVC reference software (HM 0.9) for coding of the similar test material. In addition, results for coding two sequences of the standard test set for HEVC are shown.

1 Introduction

Since natural image and video signals are acquired by a physical process, they can be regarded as degraded by noise through this acquisition process, i.e., inaccuracy of the process or physical limitations. The noise characteristics depend on the acquisition process itself. For example in digital photography the noise amount gets larger if the resolution increases using the same sensor size.

Current standards do not take into account the noise within the signal to be compressed. Especially for high quality video compression (i.e., low QP values), the current standards are not that efficient, because the noise is still present in the video at high to medium quality. It has been shown that noise removal before compression can lead to reasonable gains in compression performance [2] [3] [4], but this is prohibitive for video coding at high qualities.

In [6] we have shown that coding efficiency of lossless compression of noisy image sequences can be improved by a denoising operation applied on the reference frame, which improves the temporal prediction. We extended this idea to lossy video coding which is interesting especially for high quality compression of noisy video data. As it was supposed that quantization inside a video encoder has noise reducing capabilities, we analyzed the quantization step inside the prediction loop of a lossy transform coder [7].

2 In-loop reference frame denoising

In inter frame coding within an HEVC encoder, the predictor is generated by motion estimation and motion compensation of the previous decoded picture. This predictor is subtracted from the current picture and the difference is coded afterwards. The difference image is considered to have less energy and thus needs less bitrate for transmission. The motion compensation process is illustrated in Figure 1.

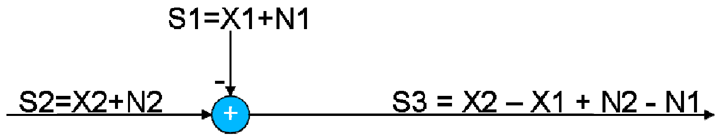

Figure 1. Motion compensation process.

In this figure, S1 describes the predictor and S2 describes the current picture which has to be encoded. If the video is assumed to be degraded by additive white noise, the current picture as well as the predictor contains noise. The difference image (i.e., the prediction error) S3 after motion compensation consists of 2 major parts, the useful signal part (X2-X1) and the noise part (N2-N1). For example, if there is no motion between two adjacent frames the useful signal part becomes zero. Usually, noise of adjacent frames is uncorrelated, thus the energy of the noise part increases.

In high quality video coding, the noise dominates the compression performance. In order to decrease the bitrate, the noise within the signal S3 has to be minimized. For high quality video coding (e.g., lossless or near-lossless coding), the noise of the current picture has also to be encoded. Thus it is not possible to remove noise directly from the difference signal S3. However, it is possible to remove noise within the predictor S1, which is also available in the decoder. Therefore, an in-loop denoising filter has been implemented in the encoder/decoder. It is worth to mention here, that low-pass filtering for interpolation of the sub-pixels removes noise by a small amount, too. Therefore, more sub-pixels are used if the noise within the video signal increases.

More information on encoding and decoding structure and the denoising algorithm can be found in JCT-VC219 [1] and in [7].

3 Results

The efficiency of the proposed algorithm for denoising of the reference frames for inter prediction has been tested using the HM 0.9 reference software. For coding tests we used the high resolution sequences (*ParkJoy*, *CrowdRun*, *InToTree*, *OldTownCross*, *DucksTakeOff*) from [9]. The sequences were cropped to a resolution of 2560x1600 pixels (from 3840x2160) and have a frame rate of 50 frames per second.

One hundred frames of each sequence were coded using the ld-he (low delay, high efficiency) settings. For the noise filtering algorithm we used a 3x3 window. The overfilter parameter has been chosen as $\xi = 6$ (see equations in [1]). The noise within the encoder and decoder has been estimated in-loop for each reference frame using the algorithm from [5]. So far the algorithm does not need additional signaling, because the whole estimation and filtering procedure is done within the encoder and decoder.

As shown in [7], the noise is still present within the reference frame for low to medium quantization parameters. For the coding experiments we used the quantization parameters $QP \in \{12,17,22,27,32,37\}$ and define a high quality (HQ) range with quantization parameters $QP \in \{12,17,22,27\}$ and a middle quality (MQ) range with quantization parameters $QP \in \{22,27,32,37\}$ which are also used for the core experiments. In the following, the rate distortion curves of the simulations for each sequence are given in Figure 2 – Figure 6.

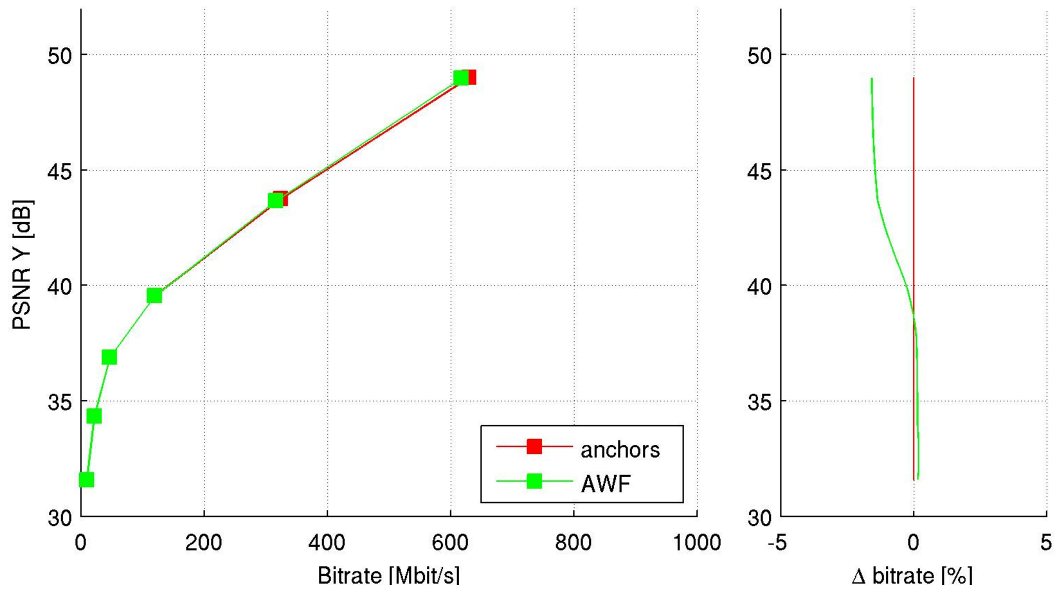

Figure 2. Simulation results for the sequence *ParkJoy*.

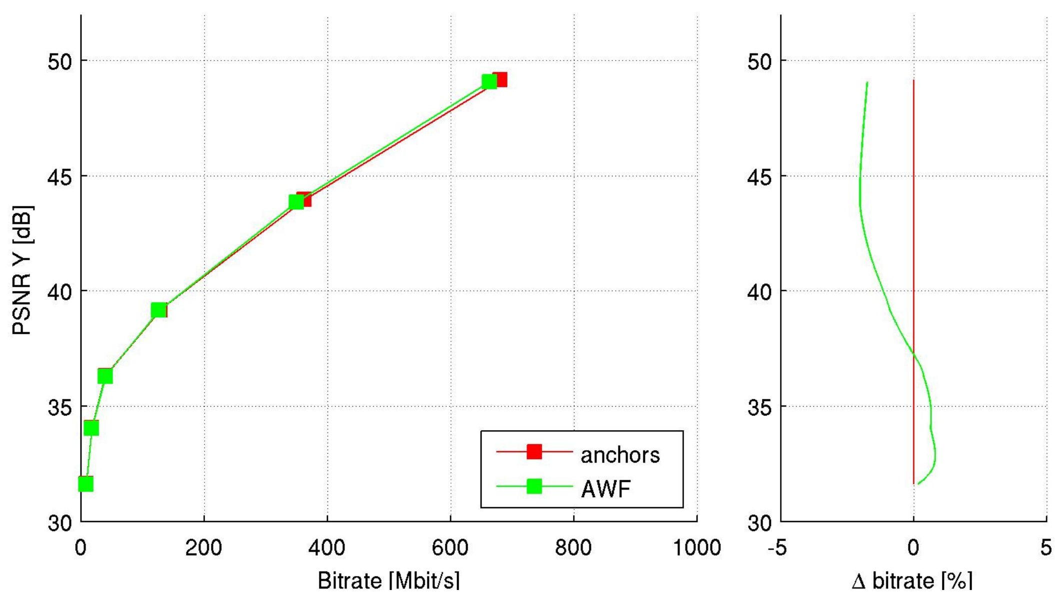

Figure 3. Simulation results for the sequence *CrowdRun*.

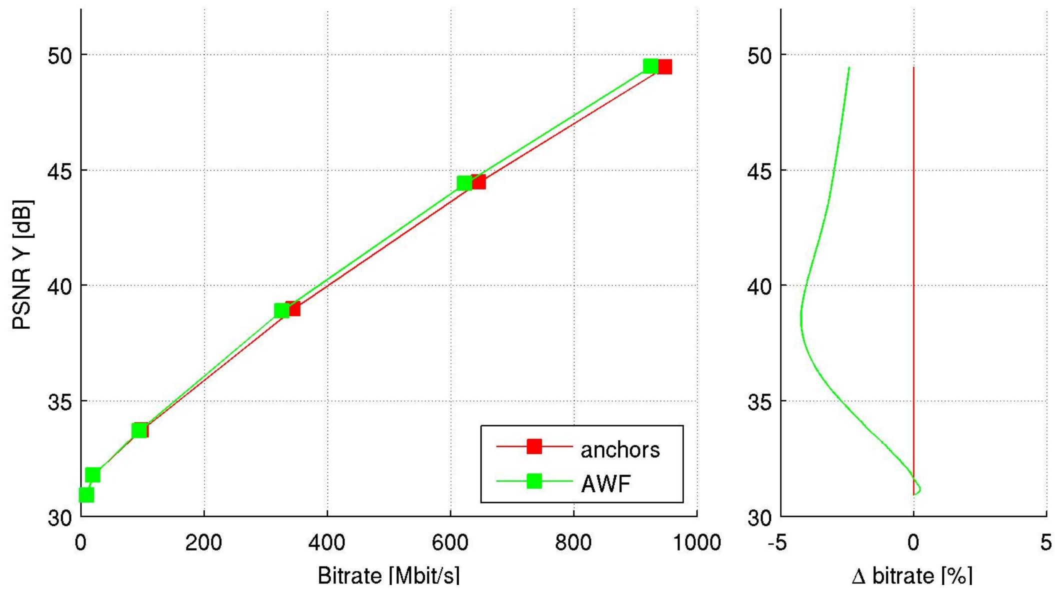

Figure 4. Simulation results for the sequence *DucksTakeOff*.

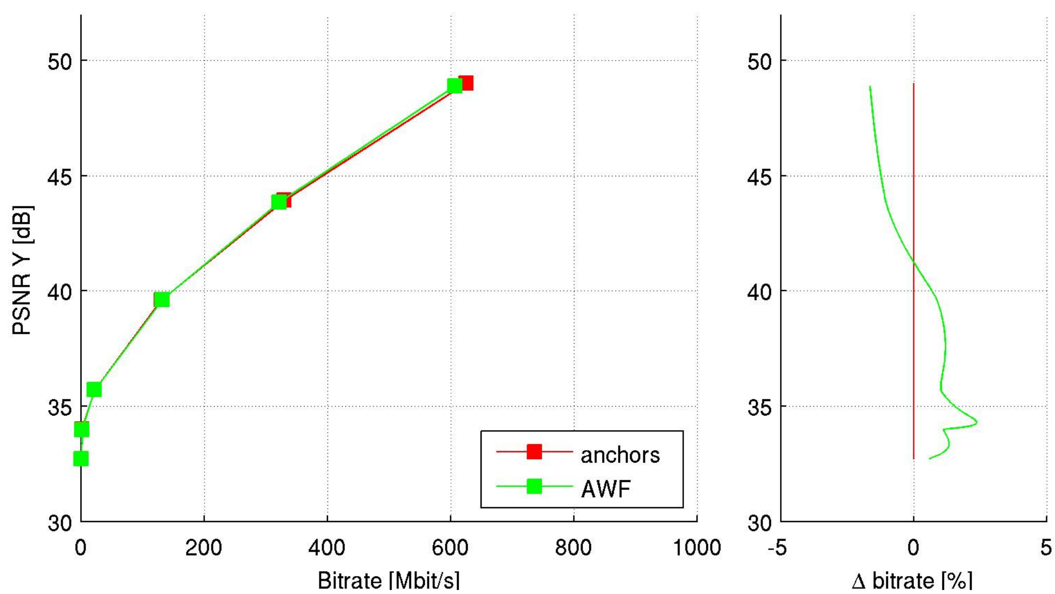

Figure 5. Simulation results for the sequence *InToTree*.

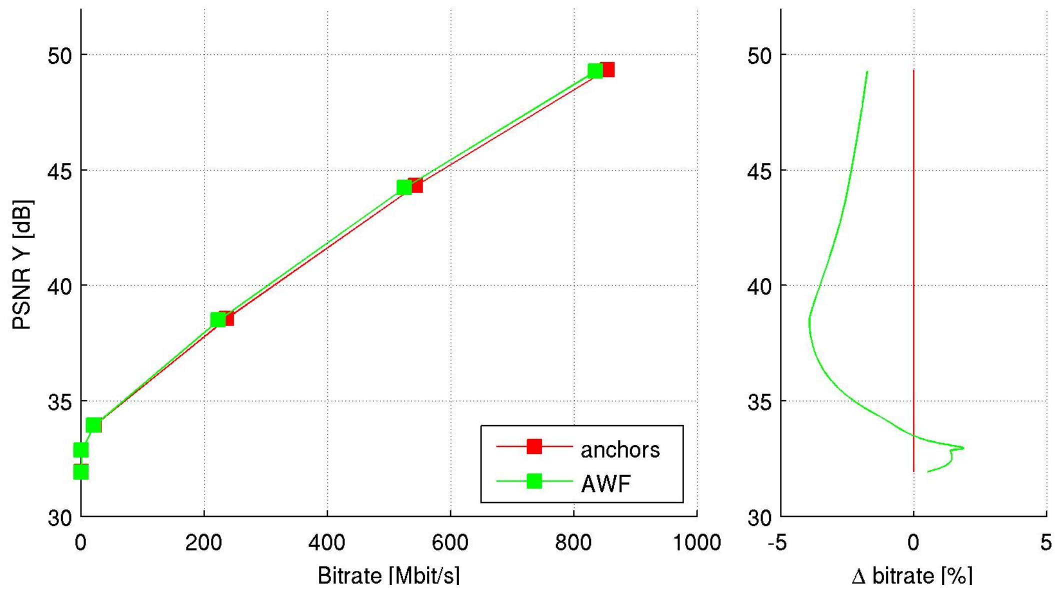

Figure 6. Simulation results for the sequence *OldTownCross.*

The figures show that for high bitrates we always have gains for these test sequences compared to the anchor results. However in low bitrate coding there is a loss in bitrate. There are two reasons for that. On the one hand, it has been shown that the noise vanishes at high quantization parameters depending on the input noise strength. On the other hand, because the noise is estimated from the quantized reference image using a low complexity algorithm, the noise estimation is not reliable for high quantization parameters where just a very small amount of noise is present. However the filtering process could be switched off adaptively for each frame which only needs one bit to be signalized by the encoder. The average bitrate differences can be observed in Table 1.

Table 1. Delta bitrate.

Sequence	delta bitrate [%] HQ	MQ	max
ParkJoy	-0.94	-0.16	-1.57
CrowdRun	-1.46	-0.55	-2.03
InToTree	-0.34	0.74	-1.63
OldTownCros	-2.60	-2.39	-3.92
DucksTakeOf	-3.27	-3.20	-4.23

It can be observed that we achieve bitrate gains on average for all tested sequences for high quality coding and we also achieve considerable gains for two of the tested sequences for medium quality coding.

Since the current HEVC test sequences do not contain considerable noise our algorithm in most cases is not effective for these sequences. We present the simulation results for the two Class B sequences *Kimono1* and *ParkScene*. As recommended, 240 frames were coded. The coding results can be observed in Figure 7 and Figure 8.

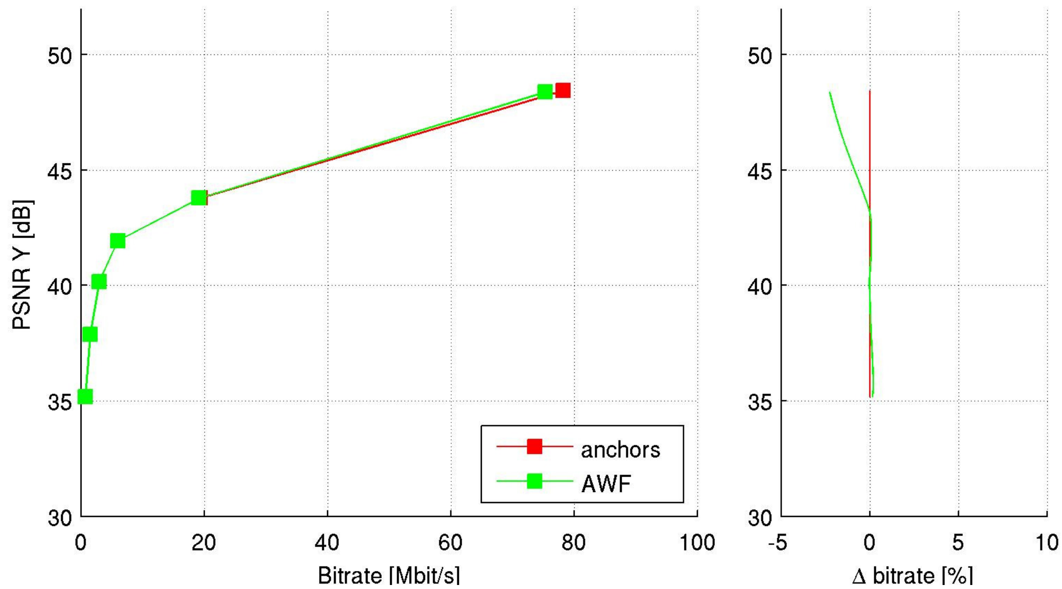

Figure 7. Simulation results for the sequence *Kimono1*.

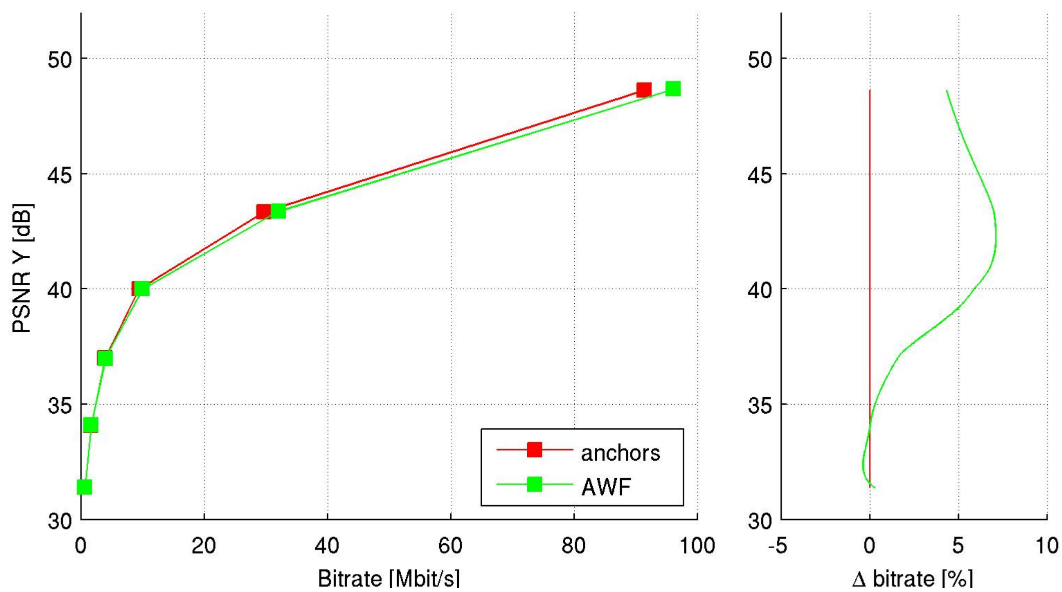

Figure 8. Simulation results for the sequence *ParkScene*.

For the *Kimono1* sequence we have gains only at very high bitrate. In case of the *ParkScene* we observed considerable loss in bitrate. This was due to false noise estimation.

4 Conclusion

In this proposal, we present an algorithm for in-loop denoising of the reference frame. The algorithm modifies the temporal predictor while the decoded picture is unchanged. The compression efficiency of the hybrid video codec can be improved for noisy image sequences when the noise within the reference frame is minimized. We implemented an MMSE estimator inside the HEVC reference software HM 0.9 in order to remove noise from the reference frame. Doing this we achieved a maximum coding gain of about 4% at high data-rates for two of the coded sequences.

5 Acknowledgement

The research leading to these results has received funding from the European Union's Seventh Framework Programme ([FP7/2007-2013]) under grant agreement ICT OPTIMIX (n° INFSO-ICT-214625).

6 References

[1] P. Amon, A. Hutter, E. Wige, and A. Kaup, "In-loop reference frame denoising," ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 WP3, document JCTVC-C219/M18259, Guangzhou, China, October 2010.

[2] H.-Y. Cheong, A.M. Tourapis, J. Llach, and J. Boyce, "Adaptive spatio-temporal filtering for video denoising," in Proceedings of International Conference on Image Processing (ICIP) 2004, Singapore, Oct. 2004, vol. 2, pp. 965-968

[3] Byung Cheol Song and Kong Wook Chun, "Motion-compensated temporal prefiltering for noise reduction in a video encoder," in Proceedings of International Conference on Image Processing (ICIP) 2004, Singapore, Oct. 2004, vol. 2, pp. 1221-1224

[4] Jim Deng, Alex Giladi, and Fernando Gómez Pancorbo, "Noise reduction prefiltering for video compression," Stanford University, 2008.

[5] John Immerkaer, "Fast noise variance estimation," IEEE Transaction on Computer Vision and Image Understanding, vol 64, no. 2, pp. 300-302, September 1996

[6] E. Wige, P. Amon, A. Hutter, and A. Kaup, "In-loop denoising of reference frames for lossless coding of noisy image sequences," in Proceedings of International Conference on Image Processing (ICIP) 2010, Hong Kong, Sept. 2010.

[7] E. Wige, G. Yammine, P. Amon, A. Hutter, and A. Kaup, "Analysis of in-loop denoising in lossy transform coding," in Proceedings of IEEE Picture Coding Symposium (PCS) 2010, Nagoya, Japan, December 2010.

[8] J.-S. Lee, "Digital image enhancement and noise filtering by use of local statistics," IEEE Transaction on Pattern Analysis and Machine Intelligence (PAMI), vol 2, no. 2, pp. 165-168, March 1980.

[9] Sveriges Television, "The SVT high definition multi format test set," ftp://vqeg.its.bldrdoc.gov/**.**

7 Patent rights declaration(s)

Siemens Aktiengesellschaft may have IPR 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).