Joint Collaborative Team on Video Coding (JCT-VC)
of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11
3rd Meeting: Guangzhou, CN, 7-15 October, 2010

Document: JCTVC-C219
WG11 Number: m18259

Title: **In-loop reference frame denoising**

Status: Input Document to JCT-VC

Purpose: Proposal

Author(s) or Contact(s):
Peter Amon, Andreas Hutter
Siemens Corporate Technology
Munich, Germany
Tel: +49 89 636 54642
Email: p.amon @ siemens.com

Eugen Wige, André Kaup
University of Erlangen-Nuremberg
Multimedia Communications and Signal Processing
Erlangen, Germany
Tel: +49 9131 85 27104
Email: wige @ lnt.de

Source: Siemens

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 inside the H.264/AVC reference software JM 15.1. It is shown that the bitrate can be decreased for (high resolution) noisy image sequences especially for higher qualities at medium to high data-rates.

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.

Noisy signals have to be compressed for effective storage and transmission. However, H.264/AVC, the state of the art in video compression, which exploits temporal correlation between adjacent frames, does not take into account the noise within the signal to be compressed. Therefore especially for high quality video compression, the H.264/AVC standard is not that efficient, because the noise is still present in the video at high to medium quality, which will be shown in later. It has been shown that noise removal before compression can lead to reasonable gains in compression performance [1] [2] [3]. However the noise reduction algorithm can damage the useful signal part significantly and therefore this scheme is not applicable to high quality video coding, where the signal should remain as unaltered as possible.

In [4] 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 [5].

2 In-loop reference frame denoising

2.1 Motivation for denoising of reference frames

In inter frame coding within an H.264/AVC video encoder or the HEVC TMuC 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.

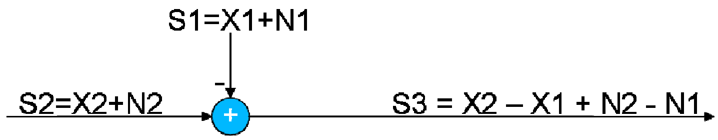

Figure 1. Motion compensation process.

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.

2.2 Block diagram

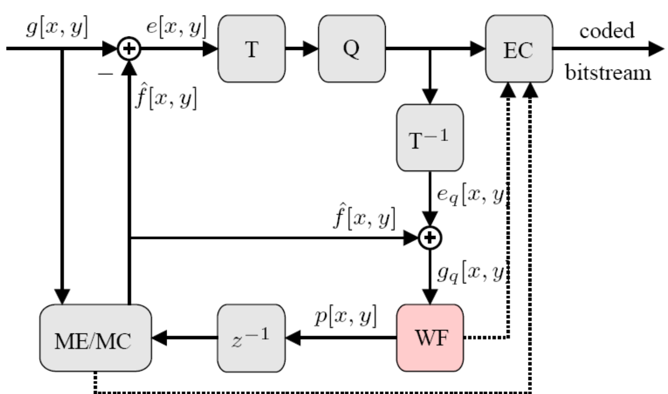

Figure 2. Encoder with in-loop denoising.

Figure 2 shows the encoder diagram with an in-loop denoising filter. In this figure, ME/MC stands for motion estimation and motion compensation, T describes the decorrelation step inside the video encoder (i.e., transform), Q describes the information reduction (i.e., quantization), and EC describes the entropy encoder within the video encoder. The WF-block (Wiener Filter) illustrates the denoising filter. We propose to remove noise from reference frames before storing them to the reference picture buffer.

The same procedure has to be done in the decoder. Each decoded picture has to be denoised before storing it in the reference picture buffer.

2.3 Denosing algorithm

For denoising of the reference frame degraded by noise a modified version of the Adaptive Wiener Filter (AWF) algorithm from [6] has been used. The algorithm is very easy to implement, does not need high computational costs, and delivered good results in lossless compression tests. We modified the algorithm by adding an additional parameter ξ. A higher ξ allows a higher noise reduction through higher averaging at the same window size, but it also introduces more errors in the useful signal part.

For the calculation of filtering parameters, the signal is considered to be stationary within a small quadratic window with side length *2M+1*. The calculation of the denoised sample is described by the following equations:

$$p(x,y) = m_f(x,y) + \frac{{\sigma_f}^2(x,y)}{{\sigma_f}^2(x,y) + \xi\sigma_{n_q}^2}(g_q(x,y) - m_f(x,y)). \quad (1)$$

$$\hat{m}_f(x,y) = \frac{1}{(2M+1)^2}\sum_{k=x-M}^{x+M}\sum_{l=x-M}^{y+M} g_q(k,l). \quad (2)$$

$$\hat{\sigma}_f^2(x,y) = \begin{cases} \hat{\sigma}_g^2(x,y) - \sigma_{n_q}^2, & if \quad \hat{\sigma}_g^2(x,y) > \sigma_{n_q}^2 \\ 0, & else \end{cases}$$
$$where \quad \hat{\sigma}_g^2(x,y) = \frac{1}{(2M+1)^2}\sum_{k=x-M}^{x+M}\sum_{l=y-M}^{y+M}(g_q(k,l) - \hat{m}_f(x,y))^2 \quad (3)$$

Equation (1) specifies the calculation of the denoised samples. For that, the local mean and the local variance are estimated as described in (2) and (3), respectively. More details on this algorithm can be found in [5].

3 Results

The efficiency of the denoising of reference frames for inter prediction was evaluated for lossy coding of noisy image sequences. The JM 15.1 version of the H.264/AVC reference software was used as a basis for the simulations. The denoising algorithm mentioned above was implemented for the in-loop-denoising of the reference frames. For the coding tests we used high resolution sequences (*ParkJoy, CrowdRun, InToTree, OldTownCross, DucksTakeOff*) from [7]. The sequences have a spatial resolution of 3840x2160 pixels, a color bit depth of 8 bits per channel (original bit depth is 16 bits per channel), and a frame rate of 50 frames per second.

Fifty frames of each of the high resolution sequences were coded. The first frame was coded as I-frame and the adjacent 49 frames were coded as P-frames, using a reference buffer size of 5 frames. For the noise filtering algorithm, we used a 3x3 window, i.e., M=1. $\sigma_{n_q}^2$ was calculated from the estimated σ_n^2.

The parameter $\xi = 3.5$ has been determined to be very good for this resolution. The estimated $\hat{\sigma}_n^2$ of the luminance channel of the unquantized sequences is shown in Table 1.

Table 1. Estimated noise $\hat{\sigma}_n^2$ within the high resolution sequences and maximum bitrate savings using in-loop-denoising in comparison to the H.264/AVC standard.

Sequence	estimated noise Variance	maximum bitrate savings [%]
ParkJoy	3.1	3.30
CrowdRun	6.5	7.70
InToTree	10.5	7.70
OldTownCross	15.9	8.10
DucksTakeOff	21.6	5.10

In Table 1 only the maximum gain is shown since the average gain largely depends on the bit-rate range used. As shown in [5], the noise is still present within the reference frame for low quantization parameters. The sequences were coded using the quantization parameters $QP \in \{10,13,16,19,22,25,28\}$. In the following, the rate distortion curves of the simulations for each sequence are given in Figure 3 - Figure 7.

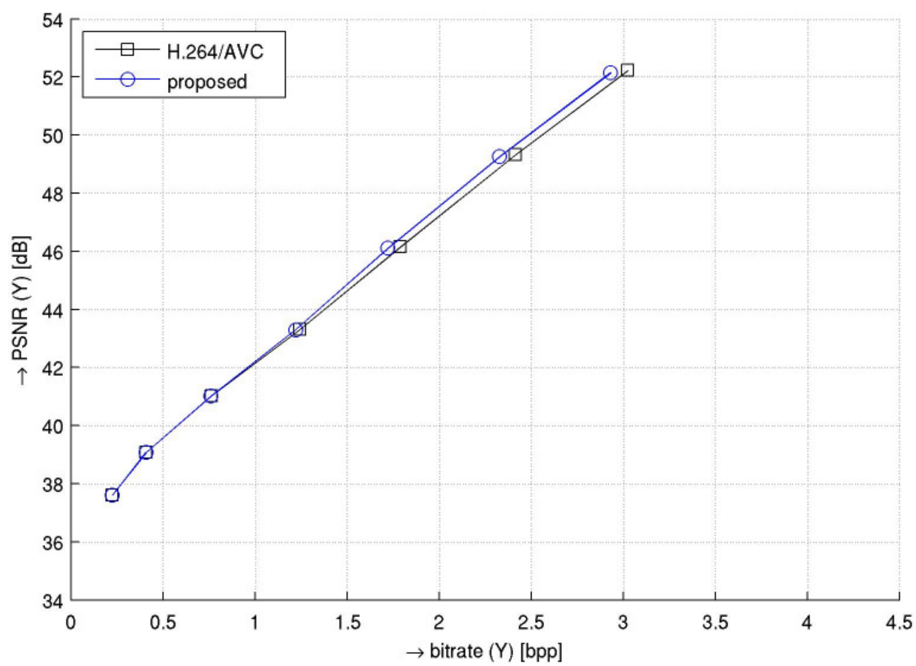

Figure 3. Simulation results for the sequence *ParkJoy*.

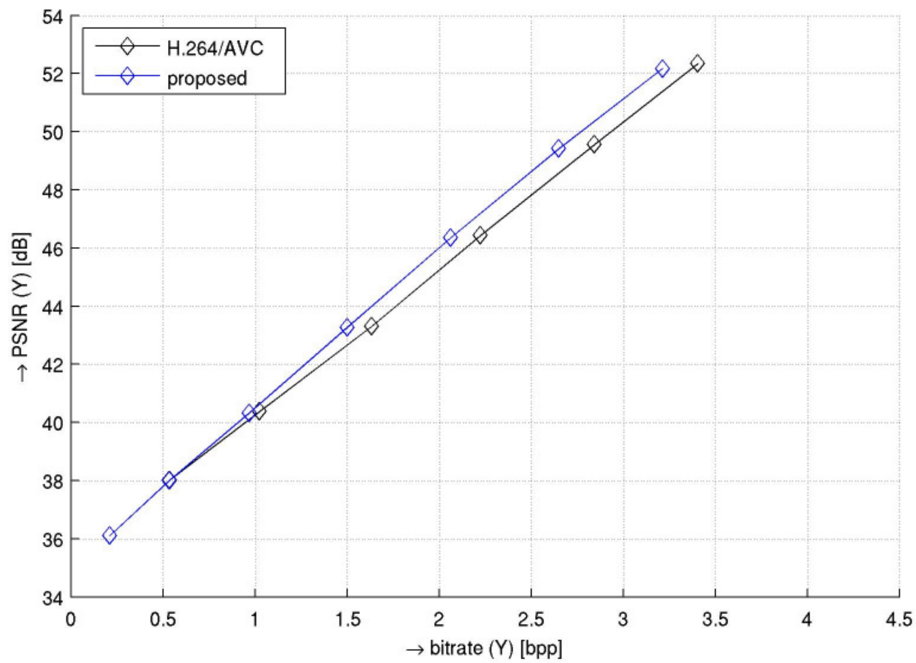

Figure 4. Simulation results for the sequence *CrowdRun*.

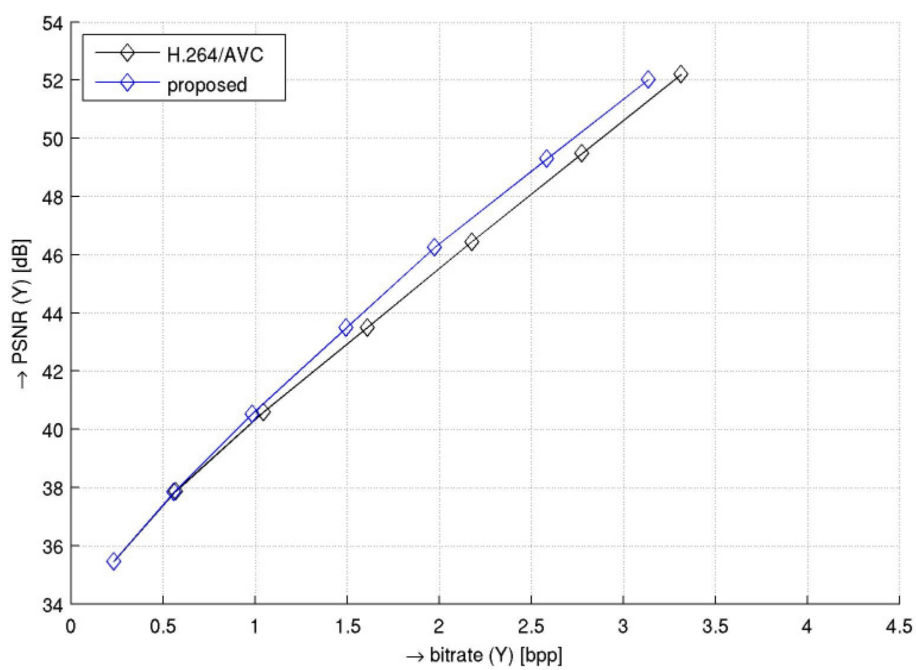

Figure 5. Simulation results for the sequence *InToTree*.

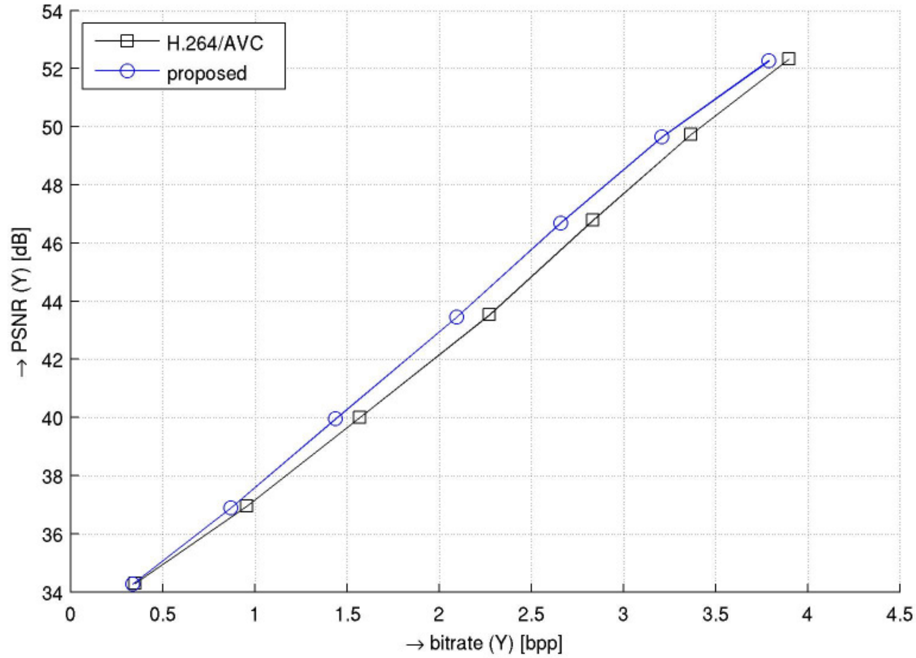

Figure 6. Simulation results for the sequence *OldTownCross.*

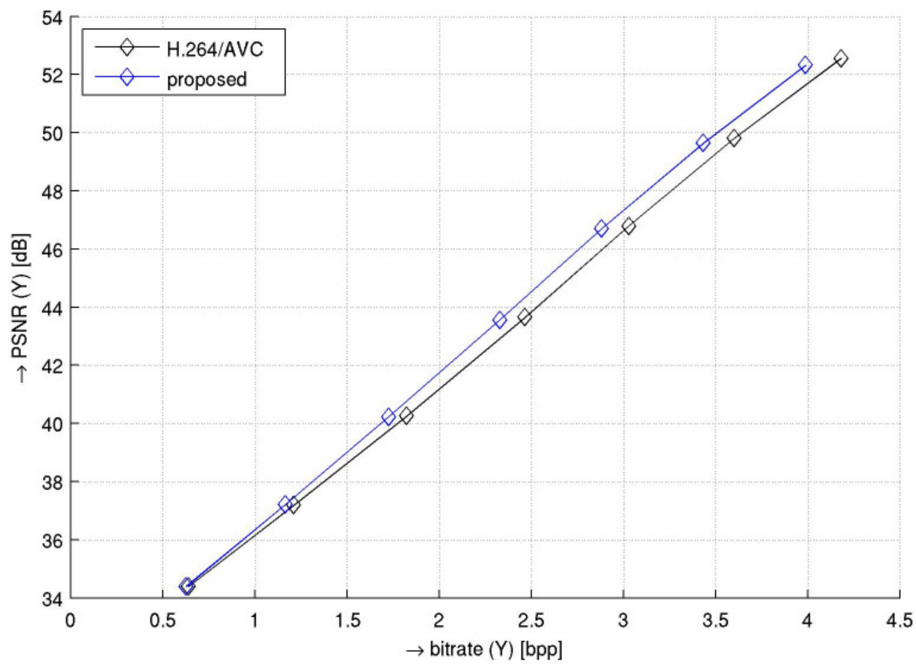

Figure 7. Simulation results for the sequence *DucksTakeOff.*

Generally, from the figures it can be seen that the distance between the two curves which are rate distortion curves for coding with JM 15.1 reference software and coding with the additional in-loop denoising filter, decreases with increasing quantization parameter. The reason for that is that the noise which is present in the video that has to be coded decreases with an increasing quantization parameter. Therefore, the proposed algorithm becomes less efficient for lower bitrate coding. Due to little noise within the *ParkJoy* sequence the coding gain for that sequence with the proposed scheme is relatively low. However, in the case of the *OldTownCross* sequence, the coding gain is much higher and is still present for $QP = 28$. This is due to the relatively high amount of noise compared to the *ParkJoy* sequence, which is still present for higher QPs, and less movement within the sequence which means that the P-modes are chosen very often. The maximum coding gains for the simulated quantization parameters are approximately 3.3%, 7.7%, 7.7%, 8.1%, and 5.1% for *ParkJoy*, *CrowdRun*, *InToTree*, *OldTownCross*, and *DucksTakeOff*. As there is a lot of movement in the *DucksTakeOff* sequence and thus P-modes are selected less often, the coding gain is not as high as in the case of the other three sequences (*CrowdRun*, *InToTree*, *OldTownCross*).

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 H.264/AVC reference software JM 15.1 in order to remove noise from the reference frame. Doing this we achieved a maximum coding gain of about 9% 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] 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

[2] 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

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

[4] 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.

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

[6] 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.

[7] 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).