Shape-Adaptive Discrete Cosine Transform for Predictor Improvement in HEVC

Formangepasste diskrete Cosinus-Transformation zur Prädiktionsverbesserung im HEVC

Eugen Wige, Andreas Heindel, André Kaup

Chair of Multimedia Communications and Signal Processing

University of Erlangen-Nuremberg

Applications with High Fidelity Requirements

- Professional applications
  - Medical applications
  - Monitoring systems
  - Industrial manufacturing
  - Digital cinema

Warner Bros, *Matrix Reloaded*

www.radiology-equipment.com, *CT Scanner*

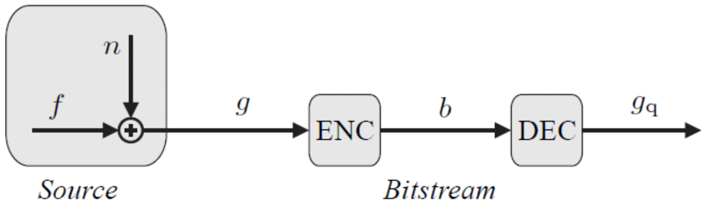
www.tagesspiegel.de, *Videoüberwachung*

Hegewald & Peschke, *Videoextensometer*

Standards for High Fidelity Compression

- Standards for professional applications
  - JPEG-LS, JPEG-2000, Motion JPEG
  - No inter-frame coding
  - Not optimal for video compression
- Standards for consumer applications
  - HEVC, H.264/AVC, MPEG-2
  - Inter-frame prediction available
  - Not optimized for high fidelity compression

- Noise within natural videos is considered as a critical issue for high fidelity compression



- Goal
  - Improve high fidelity inter-frame compression of noisy video sequences

Outline

- Motivation for in-loop denoising in high fidelity compression
- Shape-Adaptive DCT for reference frame filtering
- Coding results for medium to high quality settings
- Reference frame filtering for low bitrate compression
- Coding results for low to medium quality settings
- Conclusion and outlook

www.radiology-equipment.com, *CT Scanner*

www.tagesspiegel.de, *Videoüberwachung*

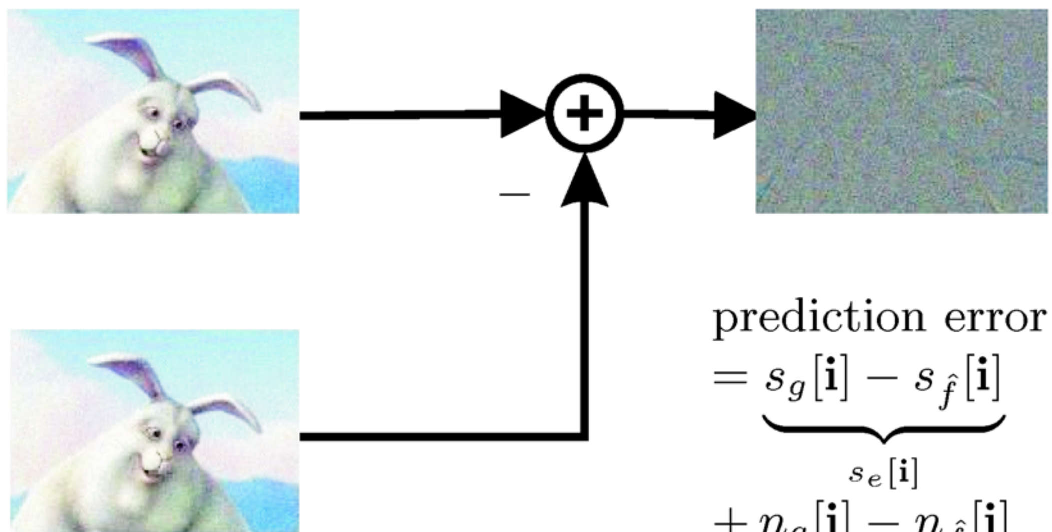
Hegewald & Peschke, *Videoextensometer*

Prediction Error Signal

- Problem
  - Prediction error signal has more noise than the current frame itself

current frame $g[\mathbf{i}]$
$= s_g[\mathbf{i}] + n_g[\mathbf{i}]$

motion compensated reference frame $\hat{f}[\mathbf{i}]$
$= s_{\hat{f}}[\mathbf{i}] + n_{\hat{f}}[\mathbf{i}]$



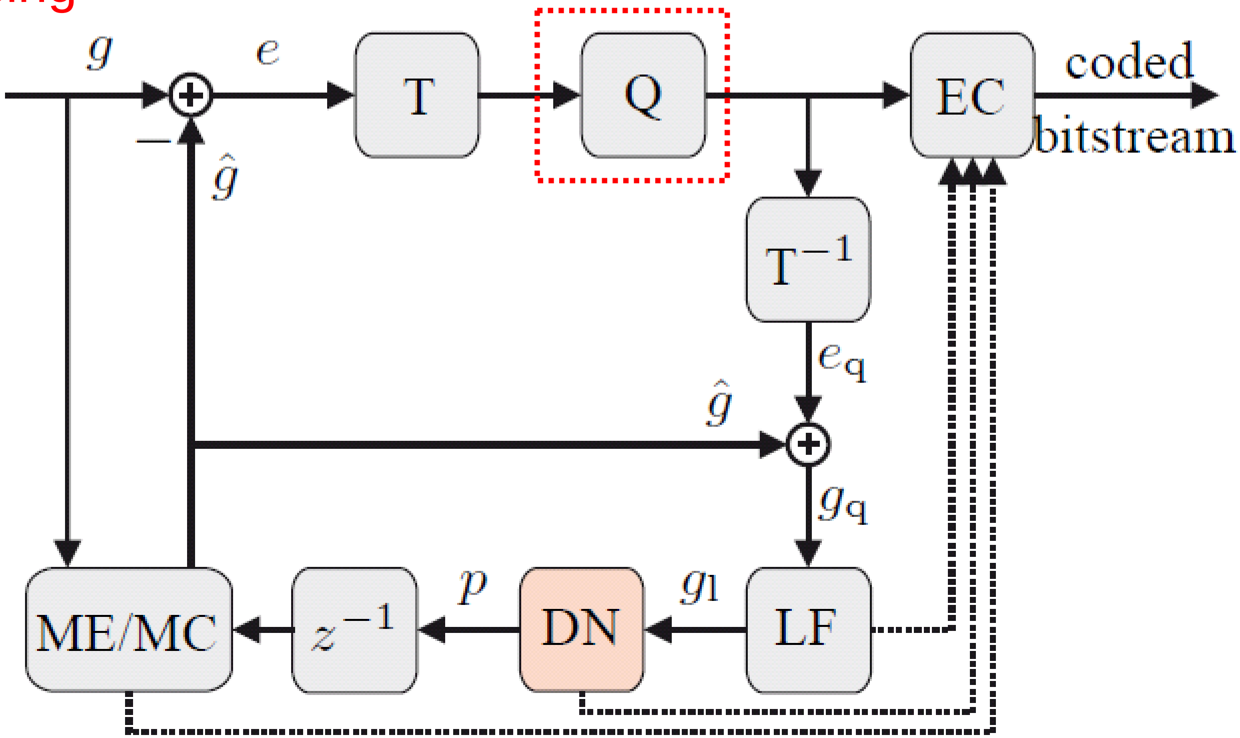
prediction error $e[\mathbf{i}]$

$$= \underbrace{s_g[\mathbf{i}] - s_{\hat{f}}[\mathbf{i}]}_{s_e[\mathbf{i}]} + \underbrace{n_g[\mathbf{i}] - n_{\hat{f}}[\mathbf{i}]}_{n_e[\mathbf{i}]}$$

- Solution
  - Remove noise from the predictor

Inter-Frame Encoder with In-Loop Denoising

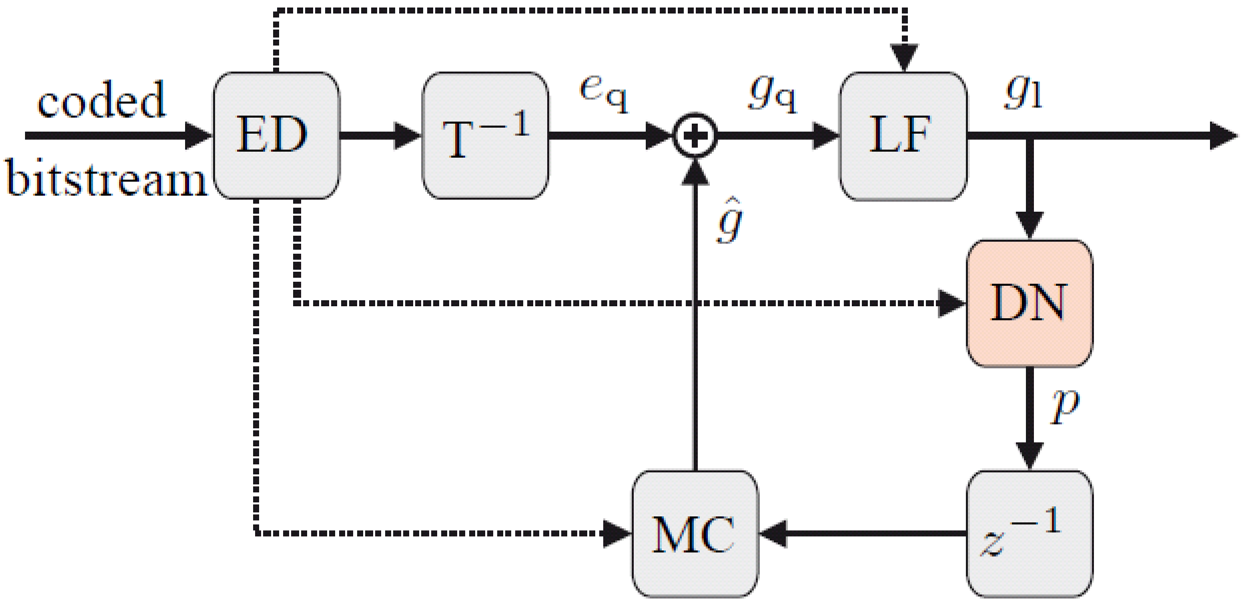
- Simplified block diagram of a lossy inter-frame encoder with in-loop denoising



E. Wige, G. Yammine, P. Amon, A. Hutter, A. Kaup, “Analysis of in-loop denoising in lossy transform coding”, PCS 2010

Inter-Frame Decoder with In-Loop Denoising

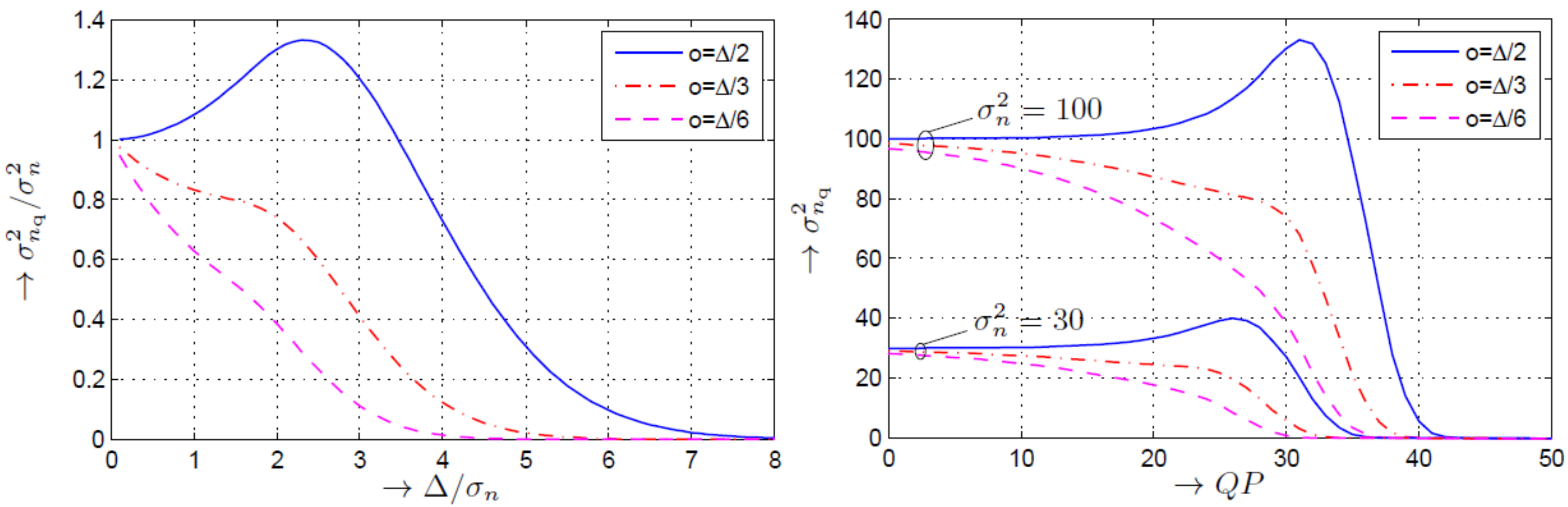
- Simplified block diagram of an inter-frame decoder with in-loop denoising



- Denoising is performed after displaying the decoded frames

Theoretical Analysis of In-Loop Quantization

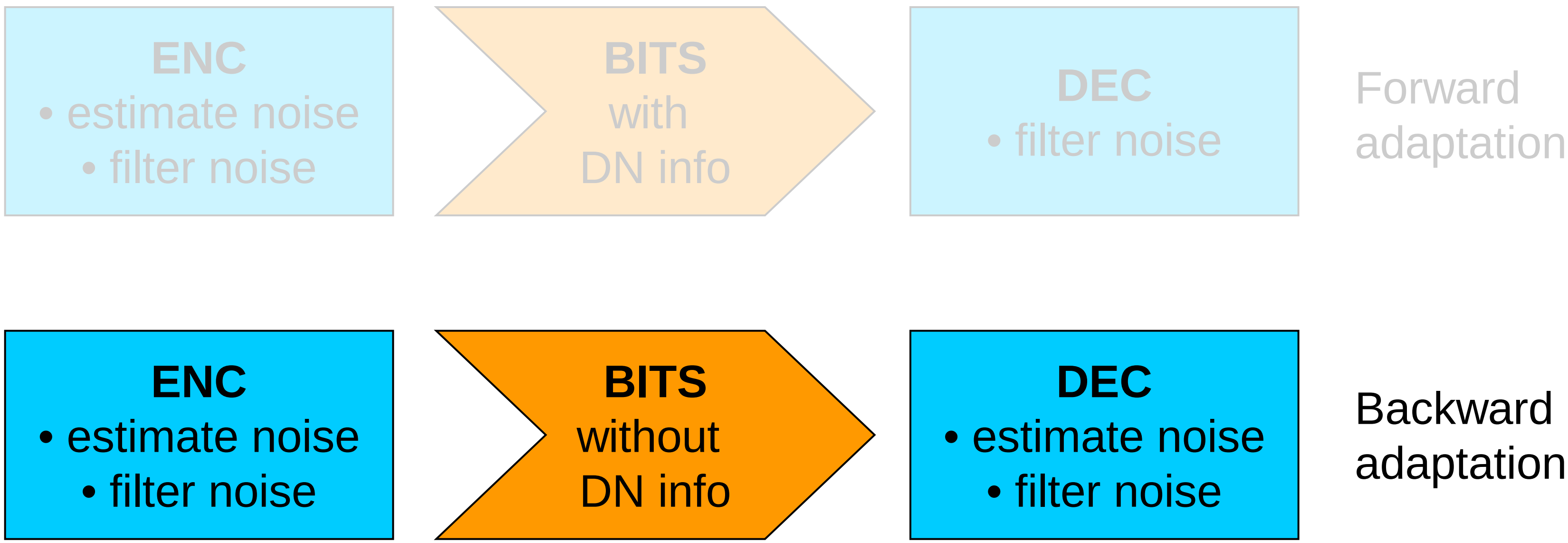
- Noise is generally decreasing for increasing QPs
- Rounding offset parameter o has additional denoising capabilities
- Noise dominates at small quantization step sizes



E. Wige, G. Yammine, P. Amon, A. Hutter, A. Kaup, "In-Loop Noise-Filtered Prediction for High Efficiency Video Coding", submitted to Transactions on Circuit and Systems for Video Technology

Realization of In-Loop Denoising

- If noise is unknown it has to be estimated
- Two ways to implement a codec with in-loop denoising



Mode Adaptive Noise Modeling

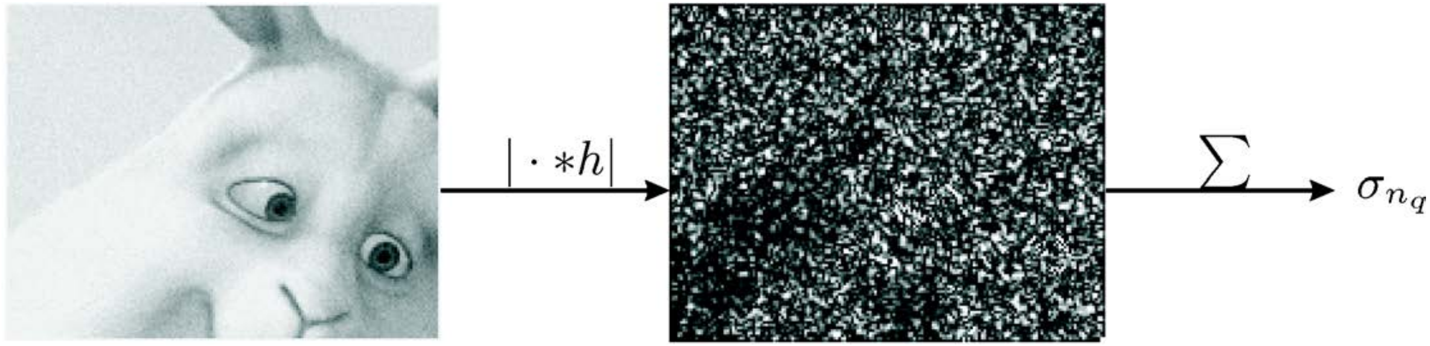
- Noise is different for different CU-types
  - Different quantization of intra and inter prediction residuals
  - Constant noise variance of skipped CUs
  - Mode Adaptive (MA) reference frame denoising
- Noise variance estimation

$$\sigma_{n_q}[c] = \frac{1}{6|\mathbf{\Omega}_c|}\sqrt{\frac{\pi}{2}}\sum_{\mathbf{j}\in\mathbf{\Omega}_c}|g_l[\mathbf{j}] * \mathbf{h}|$$

$\mathbf{h} =$

1	−2	1
−2	4	−2
1	−2	1

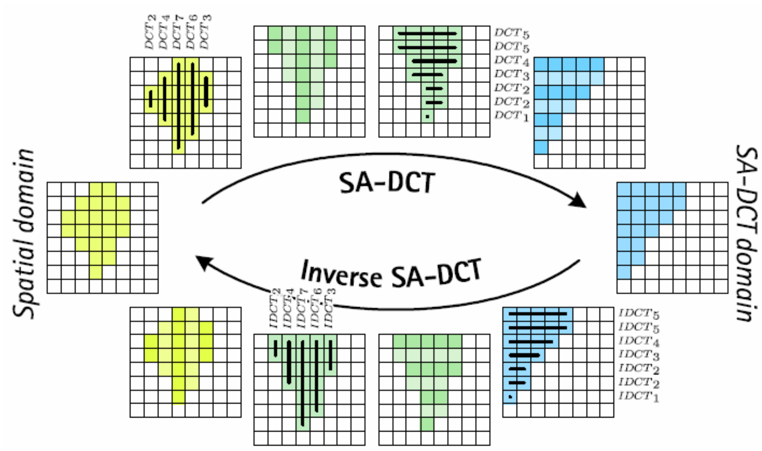
- Mode class *c={Intra, Inter, Skip}*

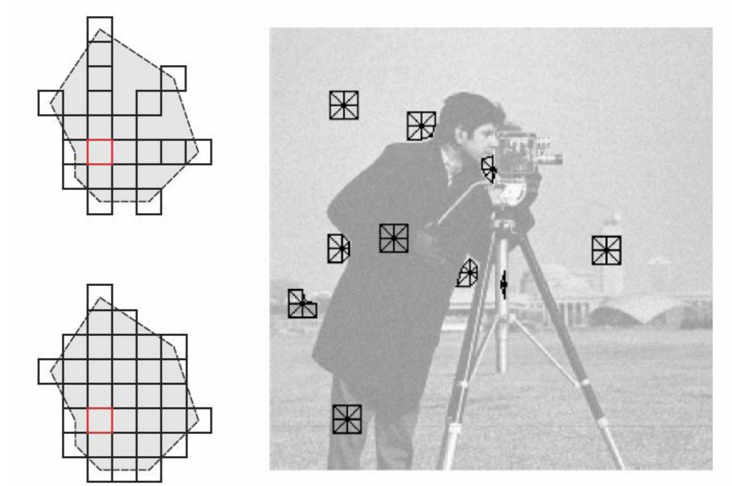


E. Wige, G. Yammine, P. Amon, A. Hutter, A. Kaup, "Mode Adaptive Reference Frame Denoising for High Fidelity Compression in HEVC", VCIP 2012

Shape-Adaptive DCT

- Introduced for coding of arbitrarily shaped regions [1]
- Supported in MPEG-4 Visual (Advanced Coding Efficiency Profile)
- Efficient denoising and deblocking tool
- Implementation of the SA-DCT algorithm according to [2]





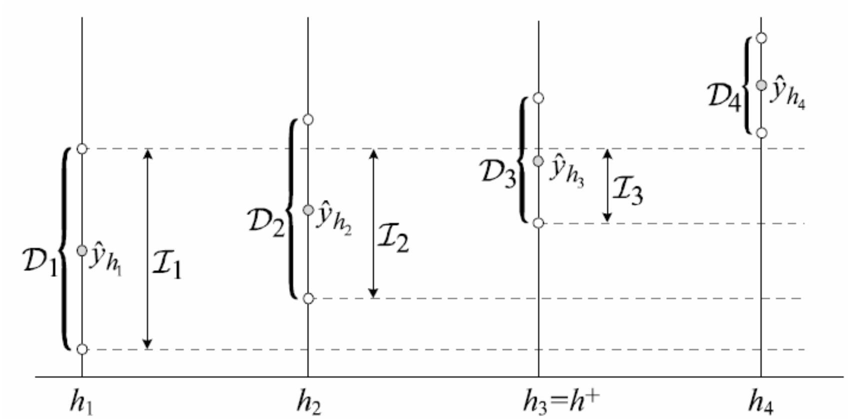
[1] T. Sikora, B. Makai, “Shape-Adaptive DCT for generic coding of video,” TCSVT 1995
[2] A. Foi, V. Katkovnik, and K. Egiazarian, “Pointwise Shape-Adaptive DCT for high-quality denoising and deblocking of grayscale and color images”, TIP 2007

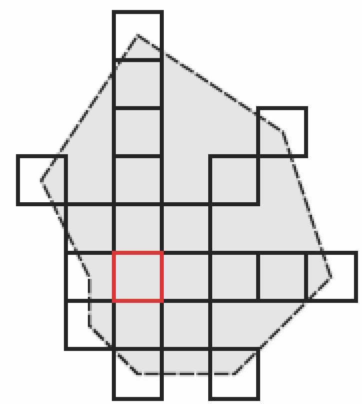
LPA-ICI-Rule for SADCT

- Local Polynomial Approximation & Intersection of Confidence Intervals
  - Search for homogenous regions (based on directional LPA-kernels)
  - Estimation of pixel value by scalar product with kernel
  - Calculation of the confidence intervals

$$\mathcal{D}_i = \left[\hat{y}_{h_i}(x) - \Gamma\sigma_{\hat{y}_{h_i}(x)},\ \hat{y}_{h_i}(x) + \Gamma\sigma_{\hat{y}_{h_i}(x)}\right]$$

  - Optimal length is given by the biggest reliable kernel





A. Foi, V. Katkovnik, and K. Egiazarian, “Pointwise Shape-Adaptive DCT for high-quality denoising and deblocking of grayscale and color images”, TIP 2007

Simulation Results

- Integration of the SA-DCT reference frame denoising algorithm into the HM-8.2 reference software
- “Lowdelay P” parameter configuration settings
- Coding of 100 frames of high and low resolution sequences
  - ClassB sequences (1920x1080)
  - ClassD sequences (416x240)
- Calculation of average bitrate savings for three quality ranges
  - Lossless compression (LL)
  - High quality (HQ, QP={12,…,27})
  - Medium quality (MQ, QP={22,…,37})

Simulation Results

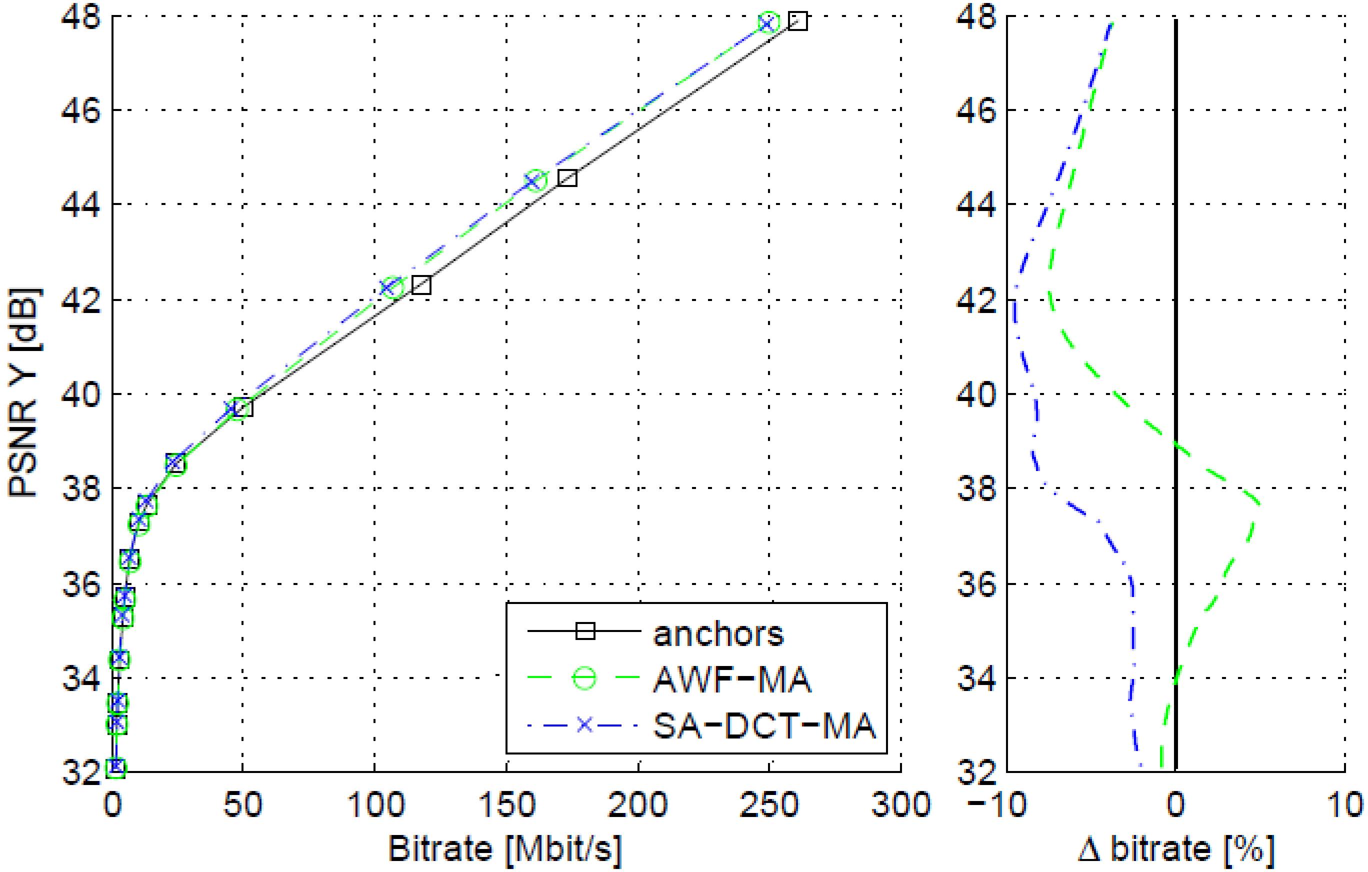
- ClassB on the top and ClassD on the bottom

Sequence	Δ bitrate in %					
	AWF-MA			SA-DCT-MA		
	LL	HQ	MQ	LL	HQ	MQ
BasketballDrive	-1.72	-3.46	0.12	-3.04	-10.23	-2.27
BQTerrace	0.03	-0.82	7.66	-1.56	-9.08	-5.21
Cactus	-1.92	-3.61	1.66	-2.61	-7.26	-3.54
Kimono1	-1.30	-1.32	-0.45	-2.25	-9.84	-3.32
ParkScene	-0.58	2.15	2.73	-1.18	1.74	-0.25
Average	**-1.10**	**-1.41**	**2.34**	**-2.13**	**-6.93**	**-2.92**
BasketballPass	3.86	3.97	0.82	-0.06	-0.16	-0.09
BlowingBubbles	1.42	4.89	1.84	-0.20	-0.48	0.02
BQSquare	9.55	27.89	12.80	-0.50	-0.01	0.65
RaceHorses	-0.07	0.07	0.14	-0.03	-0.15	-0.03
Average	**3.69**	**9.21**	**3.90**	**-0.20**	**-0.20**	**0.14**

E. Wige, G. Yammine, P. Amon, A. Hutter, A. Kaup, “Mode Adaptive Reference Frame Denoising for High Fidelity Compression in HEVC”, VCIP 2012

Simulation Results

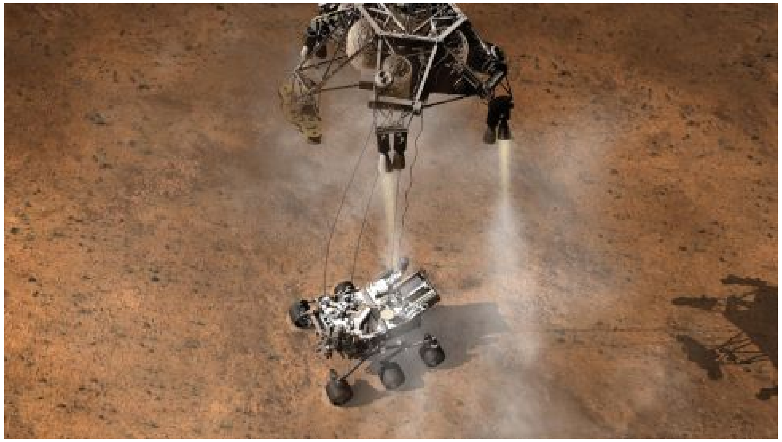
- Rate distortion and relative bitrate savings for *Cactus* (ClassB)



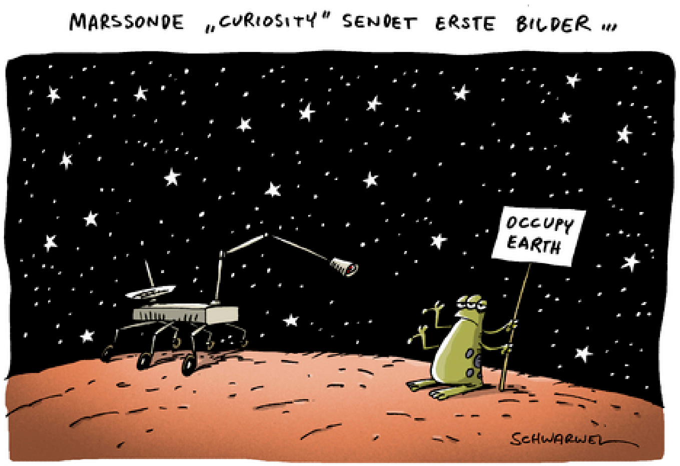
Application to Low Bitrate Coding

- Sometimes compression for very low bitrate transmission is needed
- An error signal is introduced into the reference frame due to coarse quantization
- Can the prediction be improved by reference frame filtering?

http://www.engadget.com, Skype *VOIP*



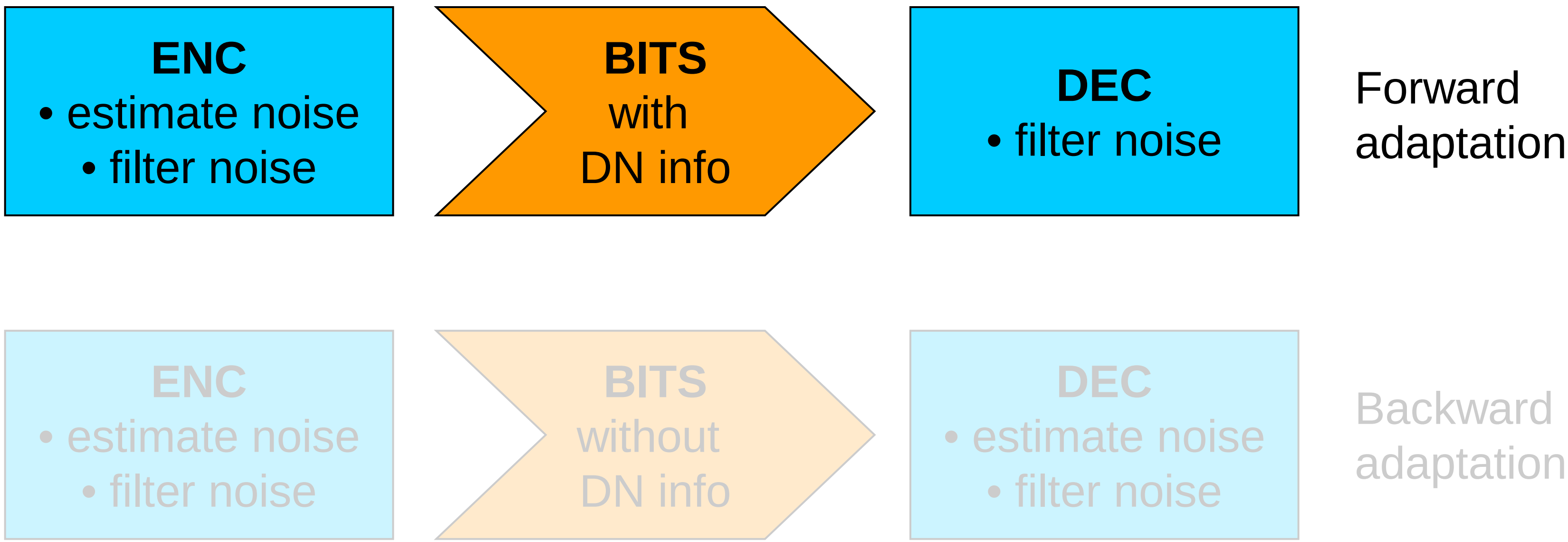
www.prosieben.de/tv/galileo, *Mars rover „Curiosity"*



de.toonpool.com, *Mars rover „Curiosity"*

Realization of In-Loop Denoising

- If noise is unknown it has to be estimated
- Two ways to implement a codec with in-loop denoising



Noise Modeling

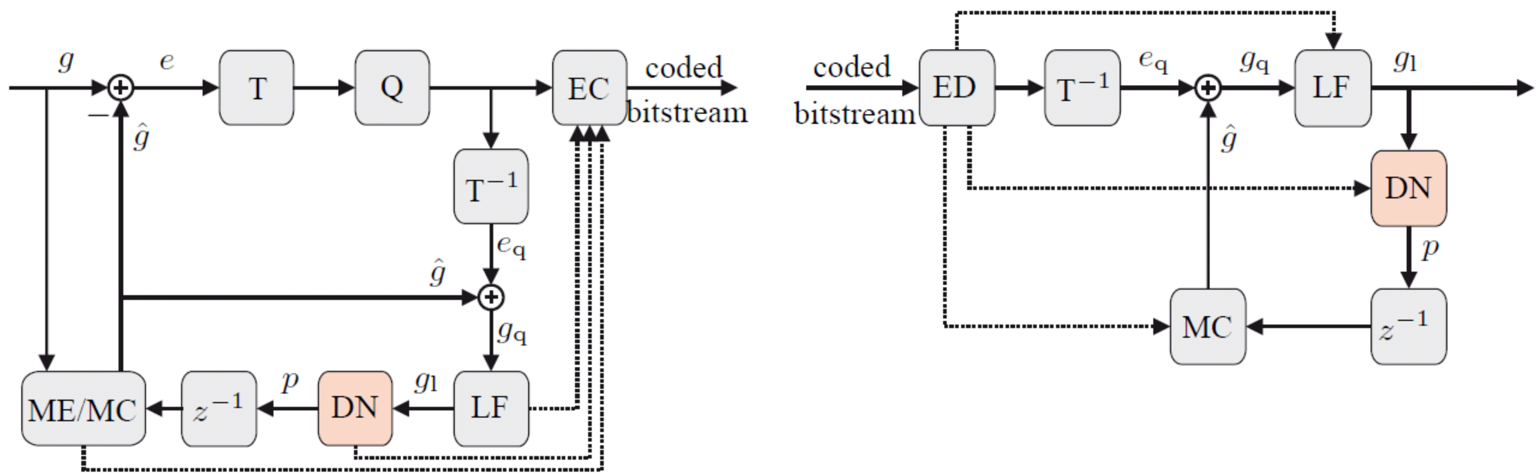
- Estimate noise standard deviation in the encoder and transmit it to the decoder
- Modeling noise variance by the MSE between the current and the reconstructed picture

$$\sigma_{\mathrm{q}}^{\mathrm{z}} = \sqrt{\frac{1}{|\Omega|}\sum_{\mathbf{j}\in\Omega}(g[\mathbf{j}] - g_l[\mathbf{j}])^2}$$

- Quantize the noise parameter for transmission

$$\sigma_{\mathrm{q}} = \min(\mathrm{round}(\sigma_{\mathrm{q}}^{z} \cdot 10)/10, 25.5)$$

- Adaptive decission between *p* and g_l



Simulation Results

- Integration of the SA-DCT reference frame filtering algorithm into the HM-8.2 reference software
- "Lowdelay P" parameter configuration settings
- Coding of 100 frames of high and low resolution sequences
  - ClassB sequences (1920x1080)
  - ClassD sequences (416x240)
- Calculation of average bitrate savings for two quality ranges
  - Medium quality (MQ, QP={22,…,37})
  - Low quality (LQ, QP={37,…,51})

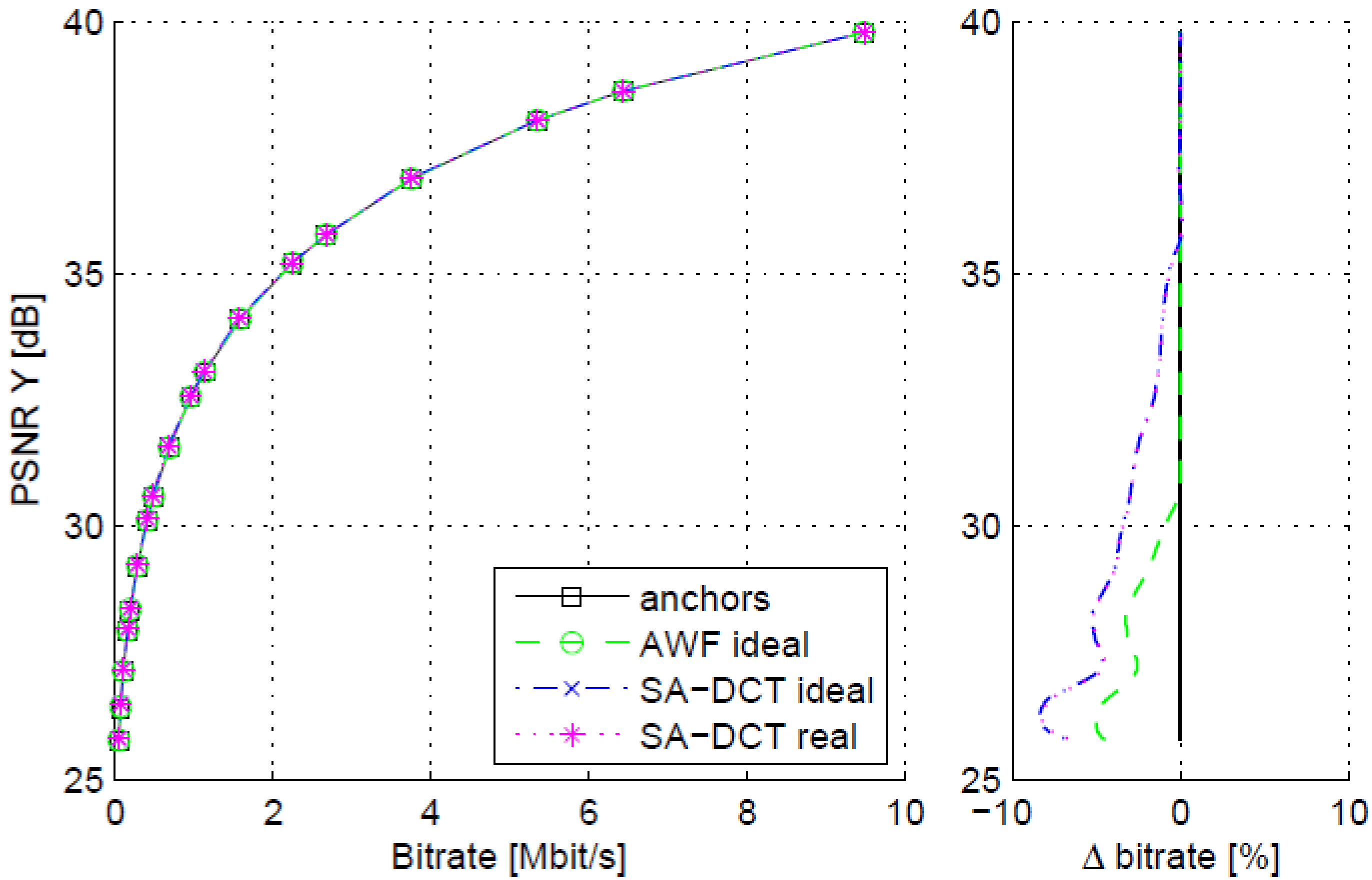
Simulation Results

- ClassB on the top and ClassD on the bottom

Sequence	Δ bitrate in %					
	AWF ideal		SA-DCT ideal		SA-DCT real	
	MQ	LQ	MQ	LQ	MQ	LQ
BasketballDrive	-0.13	-2.11	-3.57	-5.47	-3.57	-5.42
BQTerrace	0.00	-0.38	-2.75	-3.78	-2.74	-3.69
Cactus	0.00	-0.88	-2.70	-2.90	-2.70	-2.84
Kimono1	-2.65	-9.42	-8.24	-10.18	-8.23	-10.12
ParkScene	0.00	-2.24	-0.58	-4.64	-0.58	-4.57
Average	**-0.56**	**-3.01**	**-3.57**	**-5.39**	**-3.56**	**-5.33**
BasketballPass	-0.09	-2.64	-1.69	-4.54	-1.61	-3.84
BlowingBubbles	0.00	-0.42	-0.10	-1.64	-0.04	-1.01
BQSquare	0.07	0.17	-0.27	-2.25	-0.21	-1.69
RaceHorses	-0.01	-0.69	-0.92	-4.34	-0.88	-3.94
Average	**-0.01**	**-0.90**	**-0.75**	**-3.19**	**-0.69**	**-2.62**

Simulation Results

- Rate distortion and relative bitrate savings for *ParkScene* (ClassB)



Summary

- Conclusion
  - Knowledge of noise can be exploited for more efficient prediction
  - Dedicated in-loop reference frame denoising filter is proposed
  - In general impact of the source noise decreases and impact of coding noise increases with increasing QP
  - Evaluation of the SA-DCT algorithm for prediction improvement
  - Significant bitrate savings can be achieved in the state-of-the-art video coding standard for strong and weak quantization

- Outlook
  - Improve the prediction efficiency for low-bitrate coding by coding variant noise modeling
  - Context adaptive entropy coding of the noise parameters
  - Combination of the forward and backward adaptive scheme
  - Use SA-DCT for output signal reconstruction