Empowered by Innovation NEC

Description of video coding technology proposal by NEC (JCTVC-A104)

Keiichi Chono, Kenta Senzaki, Hirofumi Aoki
Junji Tajime, and Yuzo Senda

NEC Corporation

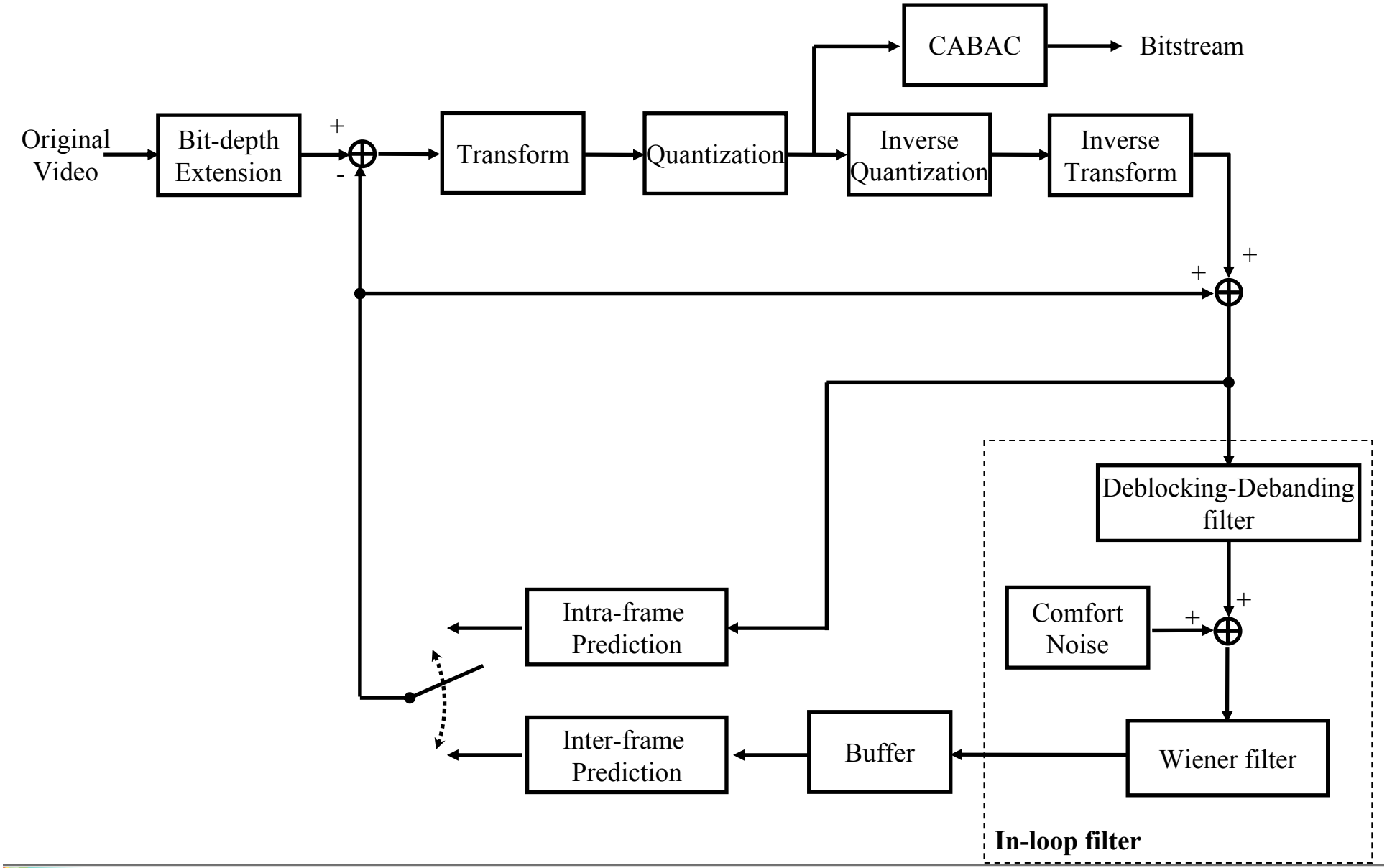
Summary

- Block-based video coding
  - Extended motion partition (32x32,32x16,16x32,16x16,16x8, 8x16, 8x8)
  - Motion vector competition
  - Variable block-size transform (16x16, 8x8, 4x4)
  - **Joint deblocking-debanding filter**
  - **Comfort noise injection**
  - Wiener filter (Frame adaptive symmetric 5x5 coefficients)
  - Internal bit-depth increase (4-bit extension for Const. 1 set configuration)

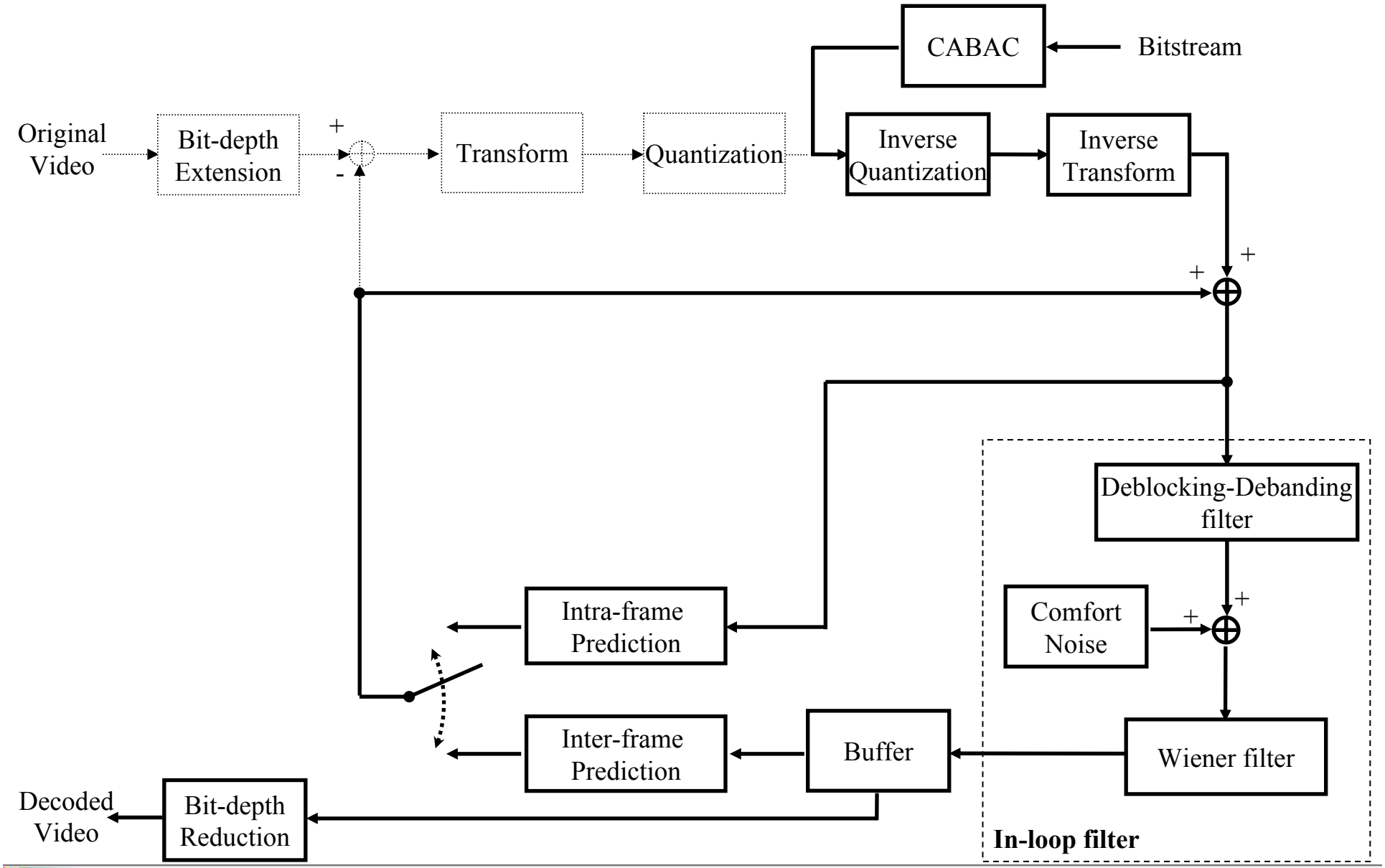
- Overall coding gain
  - Alpha anchor: 0.63 dB in BD-PSNR and -16.08 % in BD-Bitrate
  - Beta anchor: -0.10 dB in BD-PSNR and +4.58 % in BD-Bitrate
  - Gamma anchor: 0.97 dB in BD-PSNR and +22.73 % in BD-Bitrate

- Feature of proposal
  - Significant banding artifact reduction
  - Negligible impacts on Rate-Distortion performance
    - Constraint set 1: -0.006 dB in BD-PSNR and +0.158 % in BD-Bitrate
    - Constraint set 2: -0.005 dB in BD-PSNR and +0.125 % in BD-Bitrate

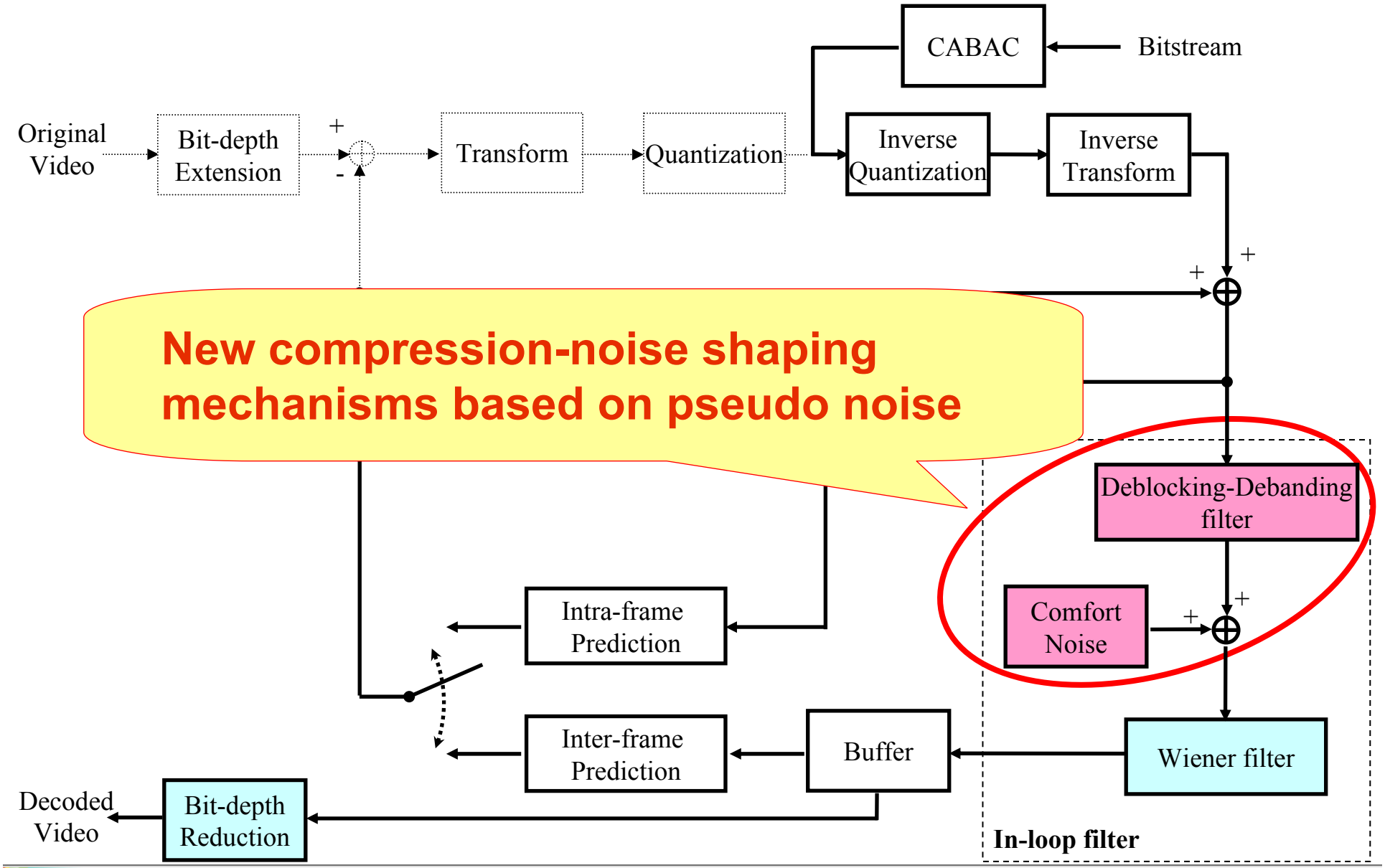
Encoder block diagram



Decoder block diagram



Decoder block diagram



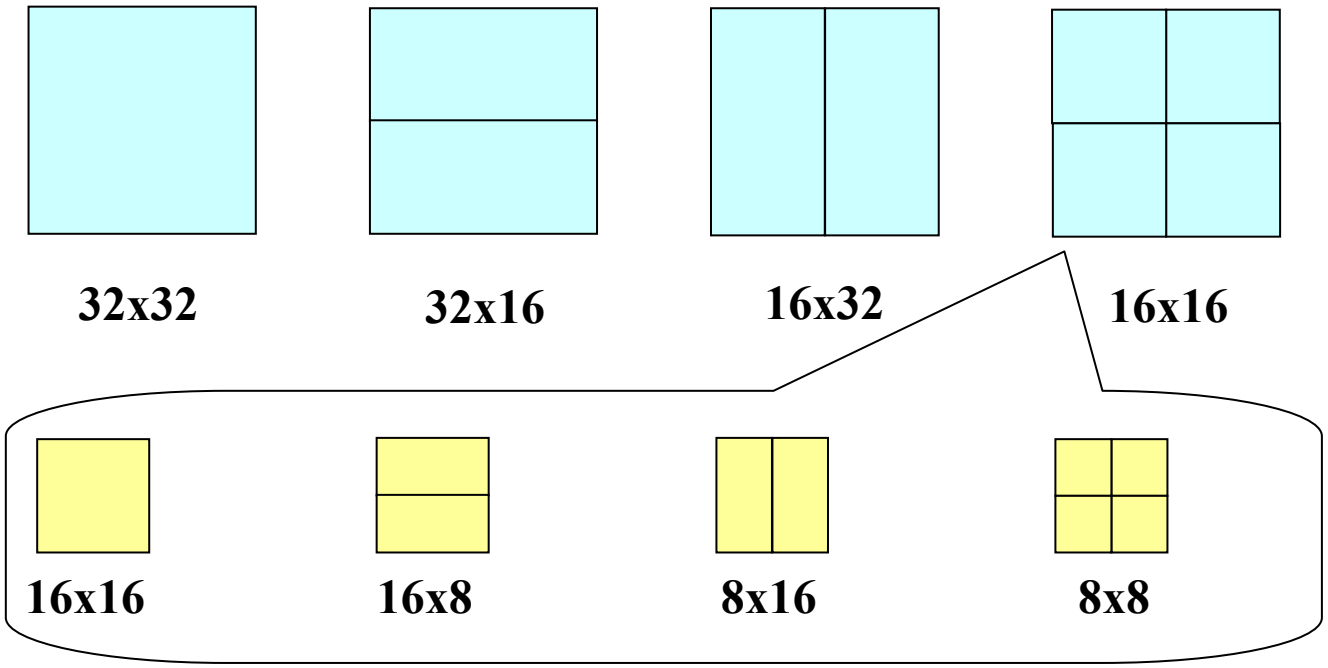
Internal Bit-Depth Increase (IBDI)

- Pixel bit-depth extension prior to encoding
  - Internal high-precision arithmetic operations
  - Quantization step size is adjusted according to the extended bit-depth
  - Pixel bit-depth reduction prior to display

- Algorithm is simple but requires higher arithmetic operations and more picture memory
  - IBDI (+4 bits) for Constraint set 1 configuration
  - No IBDI for Constraint set 2 configuration

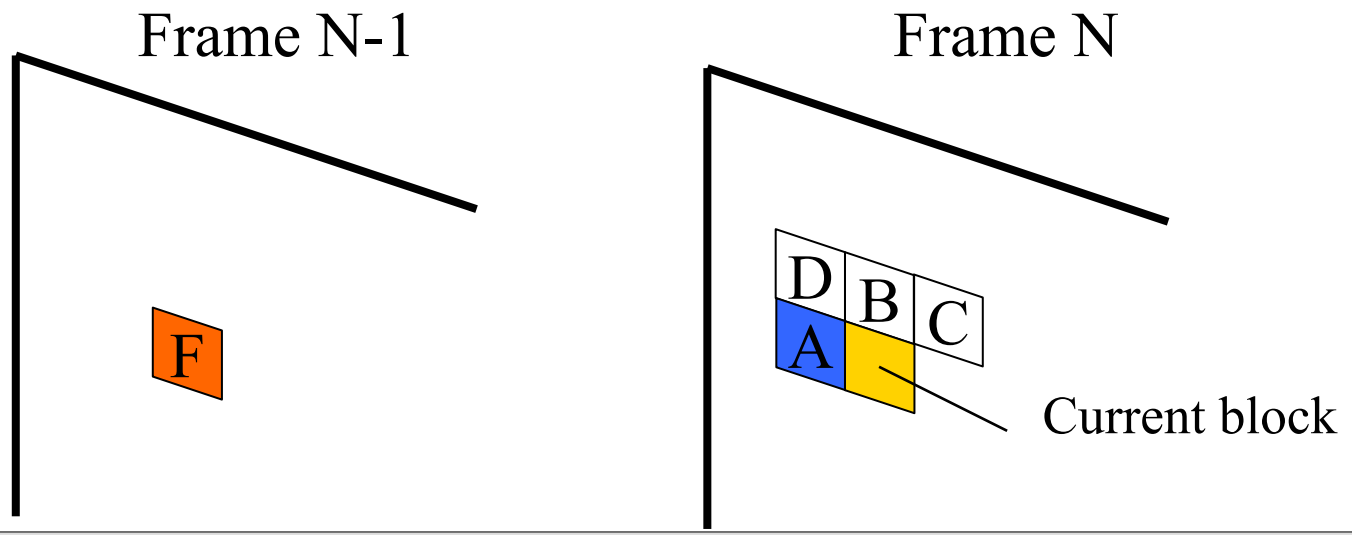
Motion partition

- Introduction of large motion partitions 32x32, 32x16, 16x32
  - Efficient MV signaling with fewer bits
- Exclusion of small partitions 8x8, 8x4, 4x8, 4x4
  - Efficient memory access for MC interpolation



Motion vector competition

- Two MV predictors
  - P-SKIP mode: Left MV (A) in addition to AVC's P-slice SKIP MV
  - P-INTER modes: Left MV (A) in addition to AVC's median MV
  - B-INTER modes: Temporal MV in addition to AVC's median MV
- Memory access and storage for MV information are similar to AVC
- One bit side information only if two predictors differ



16x16 Integer Cosine Transform (1/2)

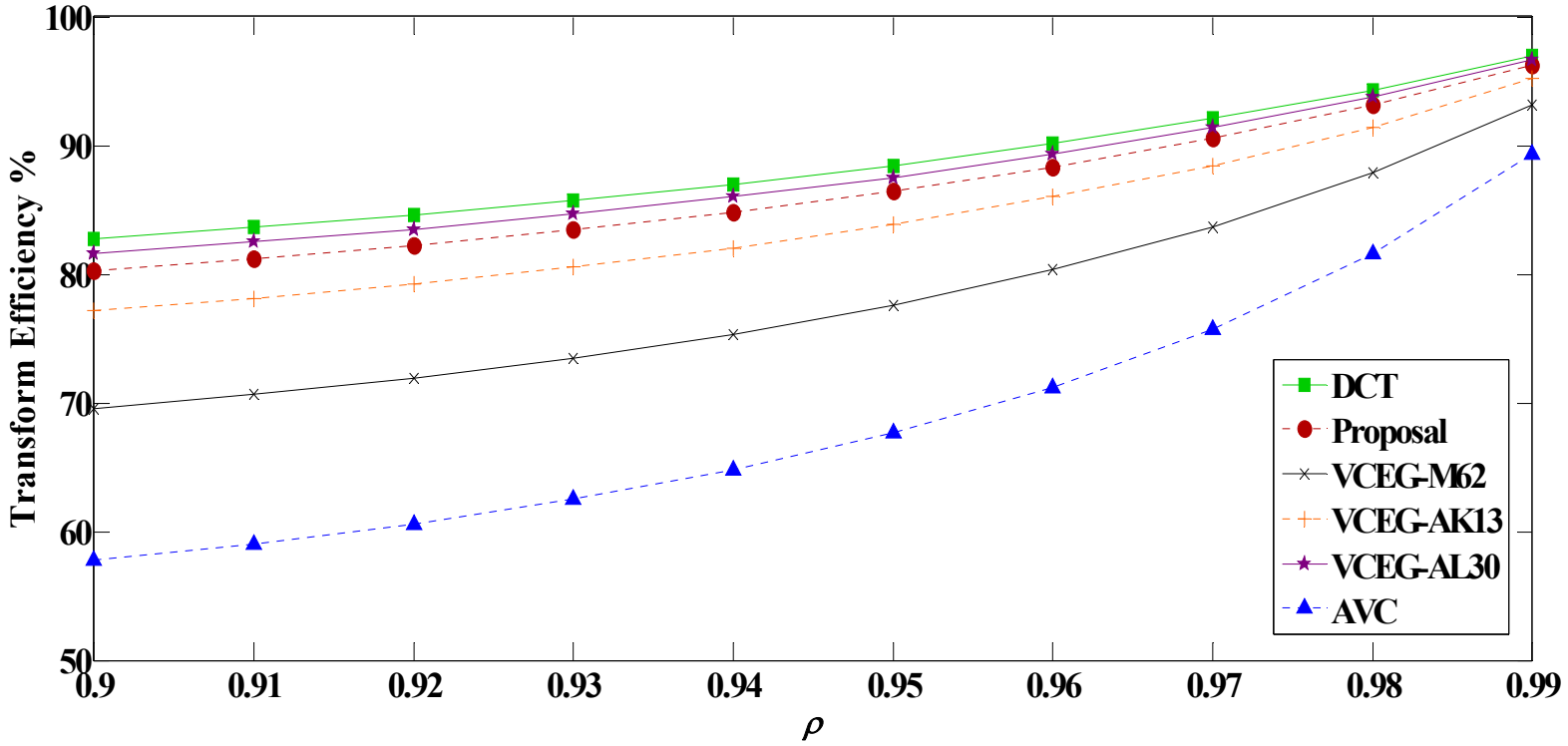
- Introduced for INTRA_16x16 prediction residual signal

- Matrix multiplication implementation but 16-bit conscious decoder inverse transform implementation

$$T_{16} = \frac{1}{32}\begin{pmatrix}
32 & 32 & 32 & 32 & 32 & 32 & 32 & 32 & 32 & 32 & 32 & 32 & 32 & 32 & 32 & 32 \\
42 & 38 & 37 & 32 & 22 & 19 & 10 & 4 & -4 & -10 & -19 & -22 & -32 & -37 & -38 & -42 \\
38 & 36 & 21 & 10 & -10 & -21 & -36 & -38 & -38 & -36 & -21 & -10 & 10 & 21 & 36 & 38 \\
38 & 22 & 4 & -19 & -37 & -42 & -32 & -10 & 10 & 32 & 42 & 37 & 19 & -4 & -22 & -38 \\
38 & 16 & -16 & -38 & -38 & -16 & 16 & 38 & 38 & 16 & -16 & -38 & -38 & -16 & 16 & 38 \\
37 & 4 & -32 & -38 & -10 & 22 & 42 & 19 & -19 & -42 & -22 & 10 & 38 & 32 & -4 & -37 \\
36 & -10 & -38 & -21 & 21 & 38 & 10 & -36 & -36 & 10 & 38 & 21 & -21 & -38 & -10 & 36 \\
32 & -19 & -38 & 4 & 42 & 10 & -37 & -22 & 22 & 37 & -10 & -42 & -4 & 38 & 19 & -32 \\
32 & -32 & -32 & 32 & 32 & -32 & -32 & 32 & 32 & -32 & -32 & 32 & 32 & -32 & -32 & 32 \\
22 & -37 & -10 & 42 & -4 & -38 & 19 & 32 & -32 & -19 & 38 & 4 & -42 & 10 & 37 & -22 \\
21 & -38 & 10 & 36 & -36 & -10 & 38 & -21 & -21 & 38 & -10 & -36 & 36 & 10 & -38 & 21 \\
19 & -42 & 22 & 10 & -38 & 32 & 4 & -37 & 37 & -4 & -32 & 38 & -10 & -22 & 42 & -19 \\
16 & -38 & 38 & -16 & -16 & 38 & -38 & 16 & 16 & -38 & 38 & -16 & -16 & 38 & -38 & 16 \\
10 & -32 & 42 & -37 & 19 & 4 & -22 & 38 & -38 & 22 & -4 & -19 & 37 & -42 & 32 & -10 \\
10 & -21 & 36 & -38 & 38 & -36 & 21 & -10 & -10 & 21 & -36 & 38 & -38 & 36 & -21 & 10 \\
4 & -10 & 19 & -22 & 32 & -37 & 38 & -42 & 42 & -38 & 37 & -32 & 22 & -19 & 10 & -4
\end{pmatrix}$$

16x16 Integer Cosine Transform (2/2)

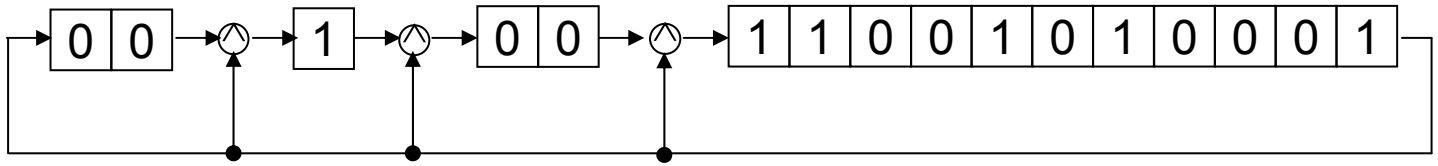
- Reasonably high de-correlation performance
  - DCT: **88.45**%
  - Proposal: **86.41**
  - AVC: **67.66**%
- Additional 68 CABAC contexts for significance_coeff_flag, last significance_coeff_flag, and coeff_abs_level_minus1



Joint deblocking-debanding filter (1/2)

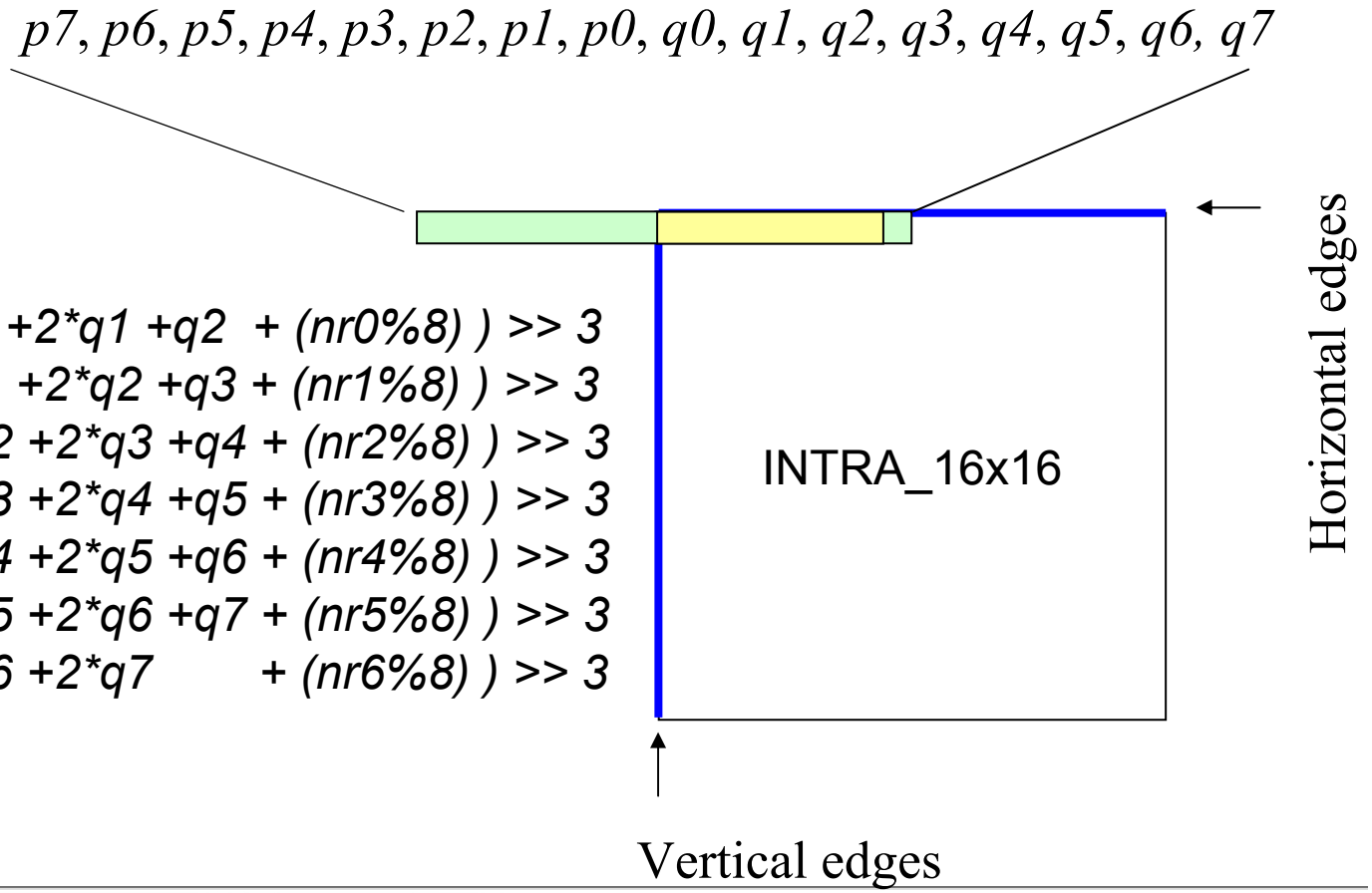
- Extension of the AVC deblocking filter, inspired by dithering filter (JVT-C056)
  - 16x16 ICT conscious dither filtering
  - Intra macroblock boundaries (BS=4) are filtered by recursive 5-tap filter with pseudo noise
  - For each edge of intra macroblock boundaries, # of pseudo-noise values, which are generated by a 16-bit Galois LFSR, is constant
  - The LFSR is initialized at each left-picture boundary macroblock to allow the easy synchronization between encoder and decoder

- Side information: The initialization value and the feedback polynomial coefficients for LFSR (32 bits)



Joint deblocking-debanding filter (2/2)

(Example) Recursive 5-tap filter with pseudo noise for vertical edges of INTRA_16x16 macroblock boundary



$$q'0 = (p1 + 2*p0 + 2*q0 + 2*q1 + q2 + (nr0\%8)) >> 3$$
$$q'1 = (p0 + 2*q'0 + 2*q1 + 2*q2 + q3 + (nr1\%8)) >> 3$$
$$q'2 = (q'0 + 2*q'1 + 2*q2 + 2*q3 + q4 + (nr2\%8)) >> 3$$
$$q'3 = (q'1 + 2*q'2 + 2*q3 + 2*q4 + q5 + (nr3\%8)) >> 3$$
$$q'4 = (q'2 + 2*q'3 + 2*q4 + 2*q5 + q6 + (nr4\%8)) >> 3$$
$$q'5 = (q'3 + 2*q'4 + 2*q5 + 2*q6 + q7 + (nr5\%8)) >> 3$$
$$q'6 = (q'4 + 2*q'5 + 3*q6 + 2*q7 + (nr6\%8)) >> 3$$

Comfort noise injection

- Noise masking inspired by Comfort Noise (SG16 Q15-B-15)
  - Enabled only when the IBDI is used
  - Pseudo noise is injected into LSBs of luma pixels on a macroblock basis
  - Wiener filter attenuates pseudo noise which is supposed to affect the MC performance in the subsequent pictures
  - The attenuated pseudo noise leads to the random dithering effects on the IBDI output images, and banding artifacts are reduced

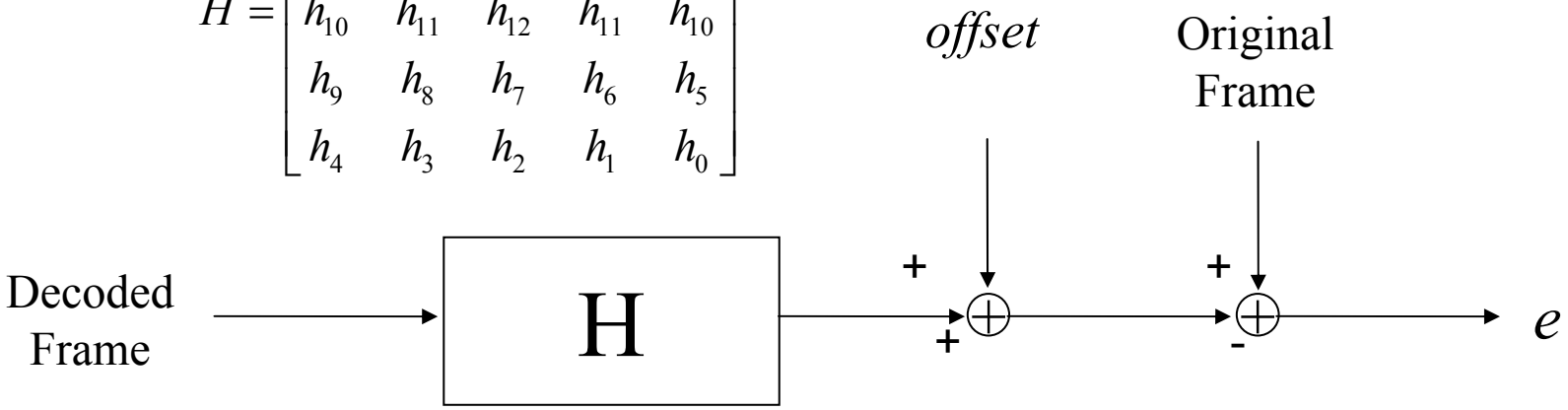
- Side information: The initialization value and the feedback polynomial coefficients for LFSR (32 bits)

Wiener filter

- Frame adaptive* symmetric 5x5 coefficients *h* (i=0,..,12) and a rounding parameter *offset*

- MMSE optimal noise reduction

$$H = \begin{bmatrix} h_0 & h_1 & h_2 & h_3 & h_4 \\ h_5 & h_6 & h_7 & h_8 & h_9 \\ h_{10} & h_{11} & h_{12} & h_{11} & h_{10} \\ h_9 & h_8 & h_7 & h_6 & h_5 \\ h_4 & h_3 & h_2 & h_1 & h_0 \end{bmatrix}$$



* Note: The combination of adaptive filter size and block adaptive filtering definitely offers better coding efficiency. Indeed, we observed some artifacts in our Constraint set 2 bitstreams of Kimono1 sequence coded at the first and second lowest rates.

Simulation results

- BD-rates and BD-PSNR relative to Anchor
  - Constraint set 1: Hierarchical B-pictures, Two reference pictures IBDI enabled, Comfort noise enabled, no multipass coding, SPS/PPS in each I-picture
  - Constraint set 2: IPPP (No hierarchical P-pictures), Two reference pictures, IBDI disabled, Comfort noise disabled

- Debanding examples

- Rate-Distortion performance changes by debanding filter and comfort noise injection
  - Constraint set 1: Comfort noise disabled, Debanding filter disabled
  - Constraint set 2: Debanding filter disabled

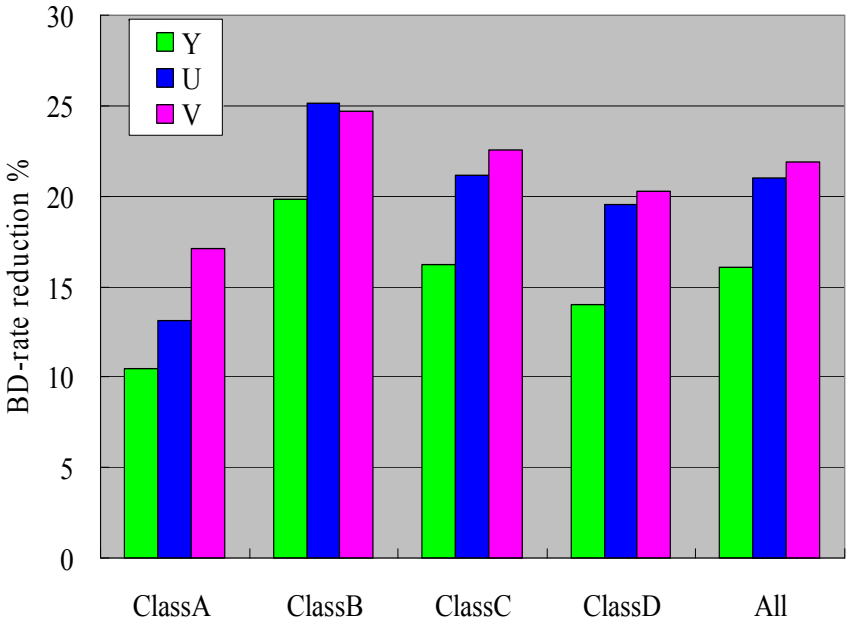
Comparison between Constraint set 1 and Alpha

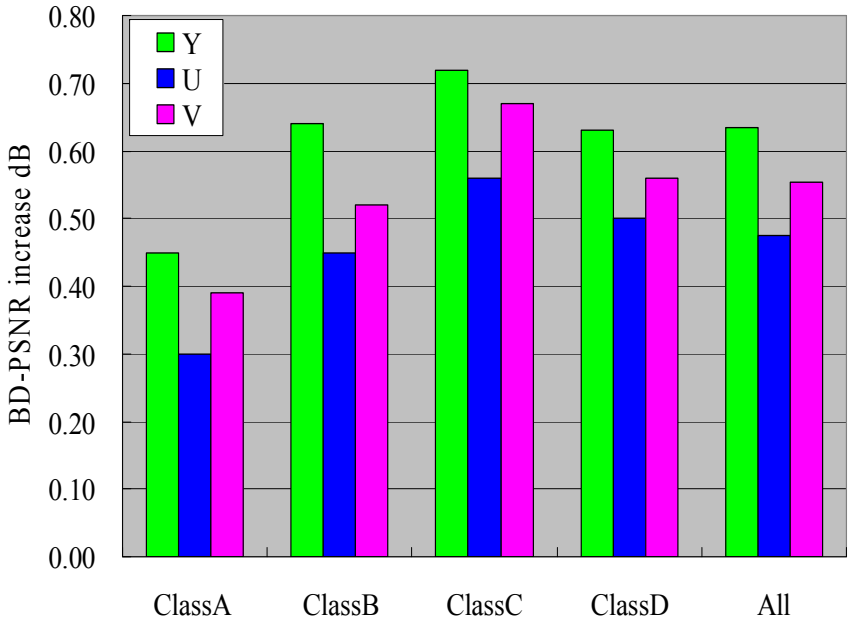
- BD-rates: Class A -10.46%, Class B -19.82%

Class C -16.23%, Class D -14.04%, and overall -16.08%

- BD-PSNRs: Class A + 0.45 dB, Class B +0.64dB

Class C + 0.72 dB, Class D 0.63dB, and overall +0.63 dB.

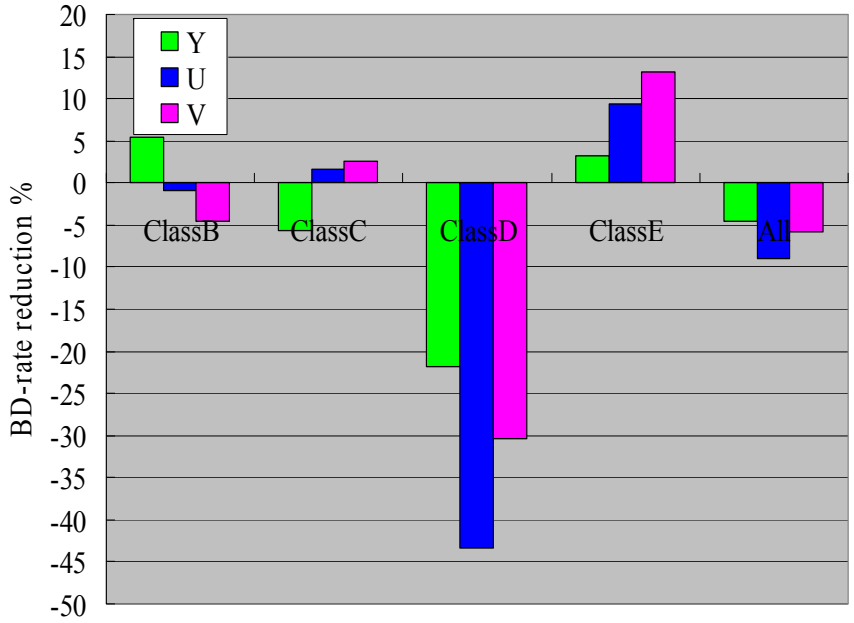


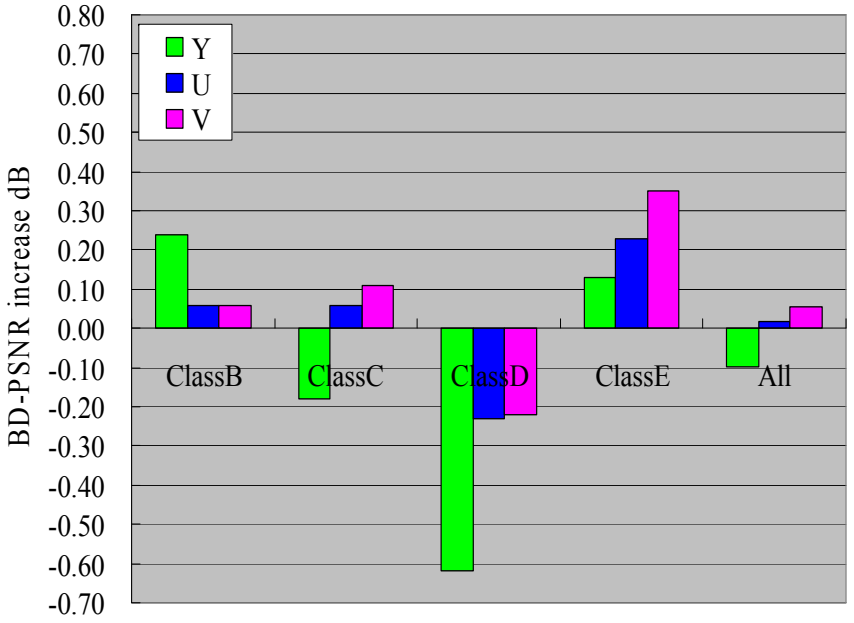


Comparison between Constraint set 1 and Beta

- BD-rates: Class B -5.49%, Class C +5.72%
Class D +21.85%, Class E -3.15%, and overall +4.58%.

- BD-PSNRs: Class B + 0.24 dB, Class C -0.18 dB
Class D -0.62 dB, Class E 0.13 dB, and overall +0.10 dB.

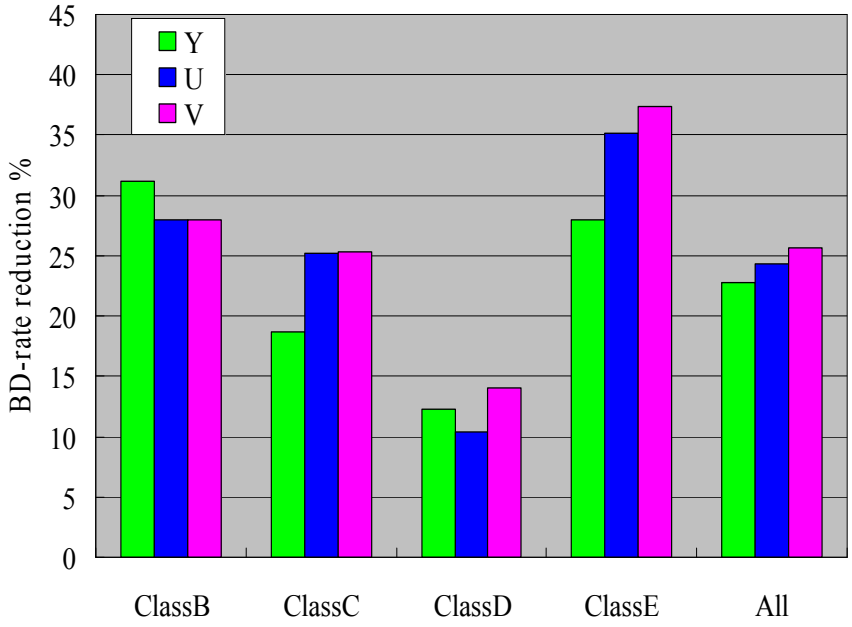


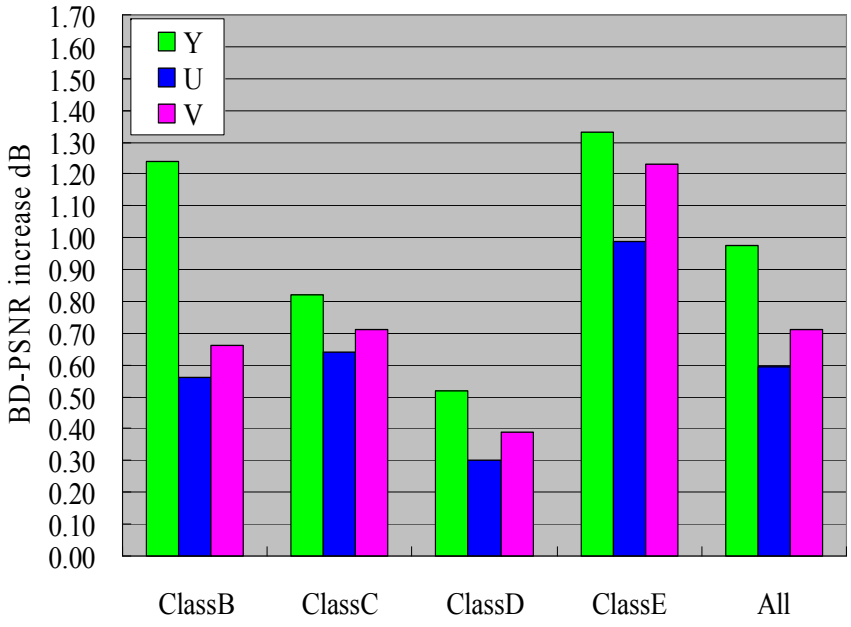


BD-rates between Constraint set 2 and Gamma

- BD-rates: Class B -31.18 %, Class C -18.71 %, Class D -12.25 %, Class E -27.98 %, and overall -22.73 %.

- BD-PSNRs: Class B +1.24 dB, Class C +0.82 dB, Class D +0.52 dB,, Class E +1.33 dB, and overall +0.97 dB.





Debanding example: Kimono1 144th frame (1/6)

Original image

Debanding example: Kimono1 144th frame (2/6)

Cropped and zoomed original image

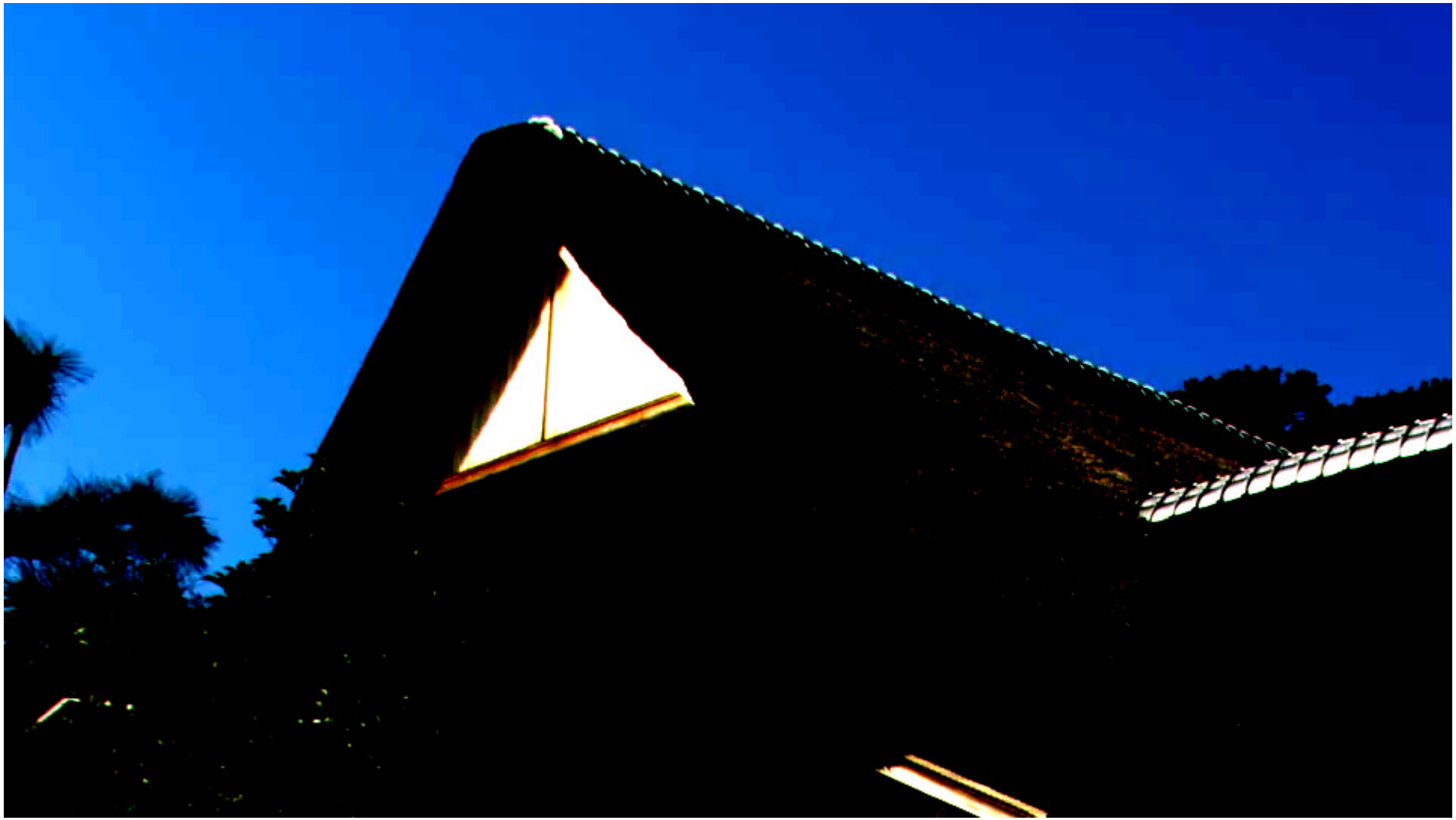
Debanding example: Kimono1 144th frame (3/6)

(Color enhanced) Cropped and zoomed original image

Debanding example: Kimono1 144th frame (4/6)

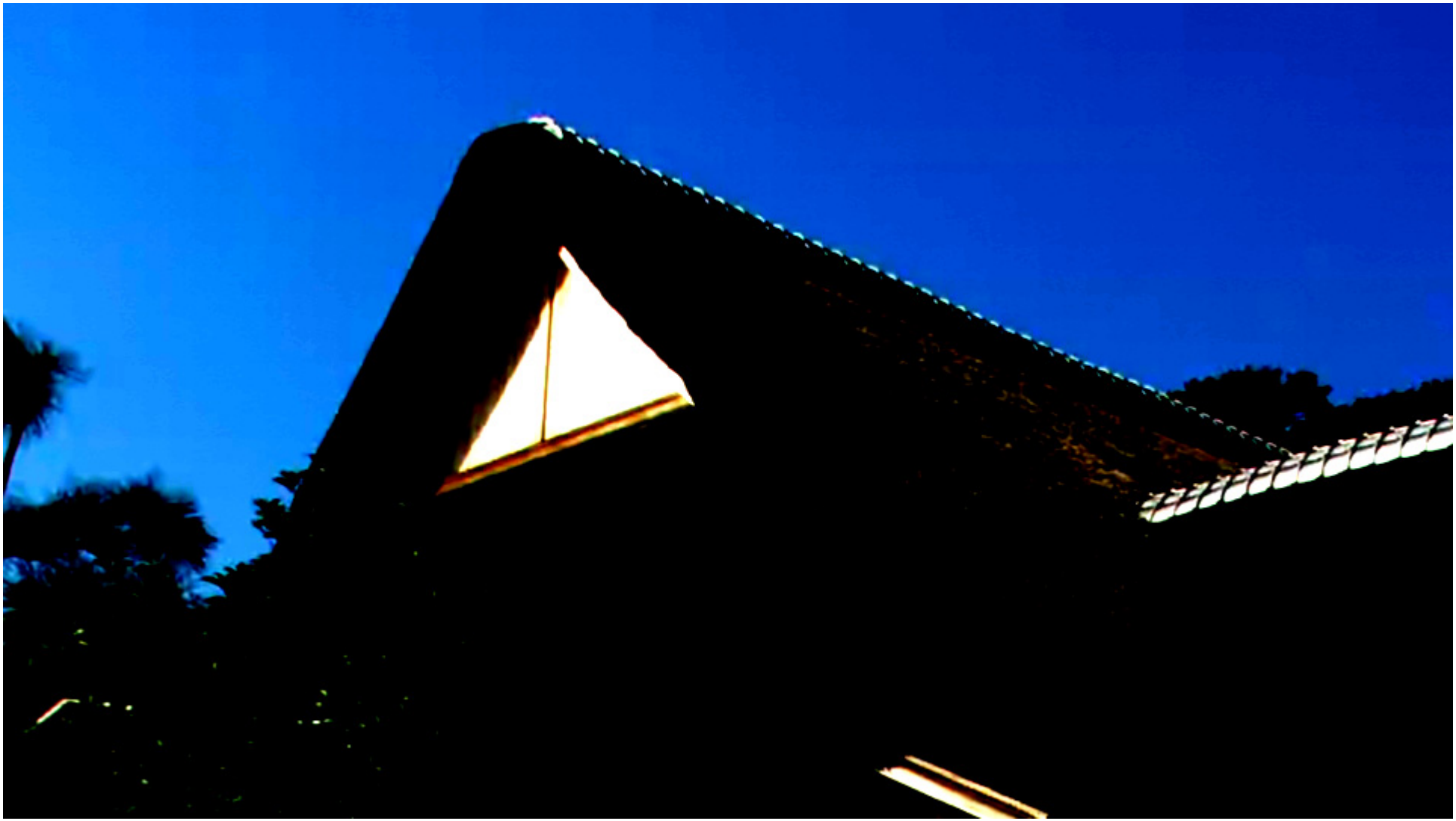
(Color enhanced) Anchor compressed 1.6Mbps

Debanding example: Kimono1 144th frame (5/6)



(Color enhanced) Anchor compressed 6.0Mbps

Debanding example: Kimono1 144th frame (6/6)



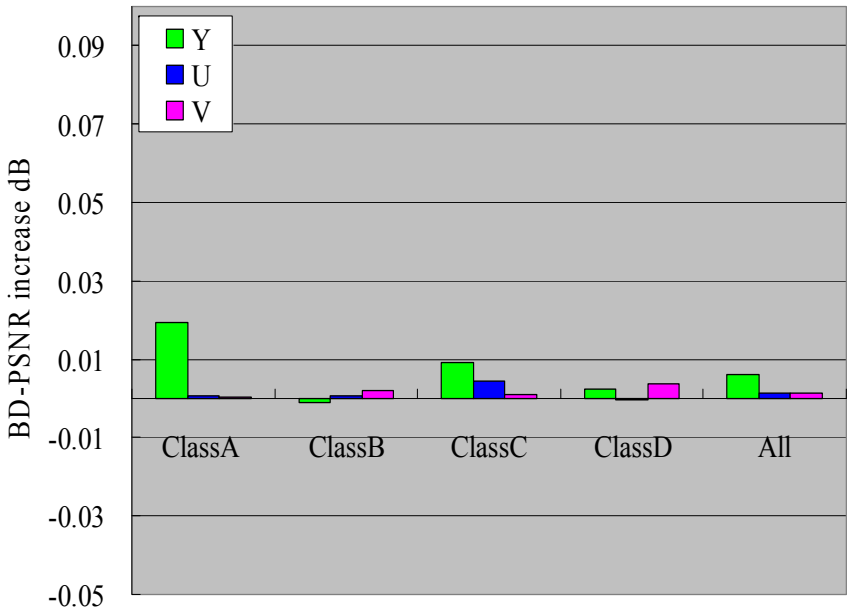
Submission compressed 1.6Mbps

Impacts on BD-PSNRs

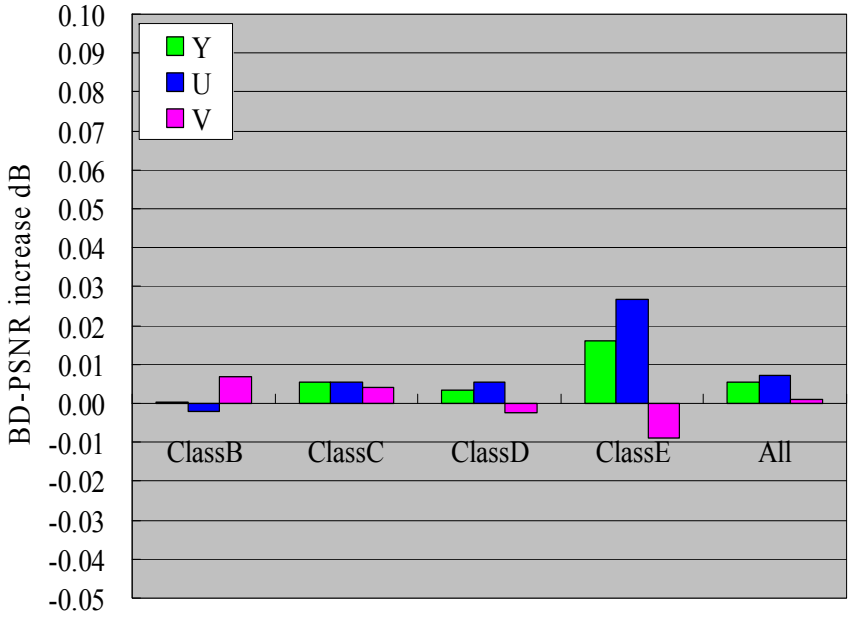
- Constraint set 1: Class A -0.020 dB, Class B 0.001 dB, Class C -0.009 dB, Class D -0.003 dB, and overall -0.006 dB.

- Constraint set 2: Class B 0.000 dB, Class C -0.005 dB, Class D -0.003 dB, Class D -0.016 dB, and overall -0.005 dB.

Constraint set 1 results



Constraint set 2 results



Conclusions

- Video coding technology based on a new in-loop filter that integrates noise-shaping and noise-reduction mechanisms

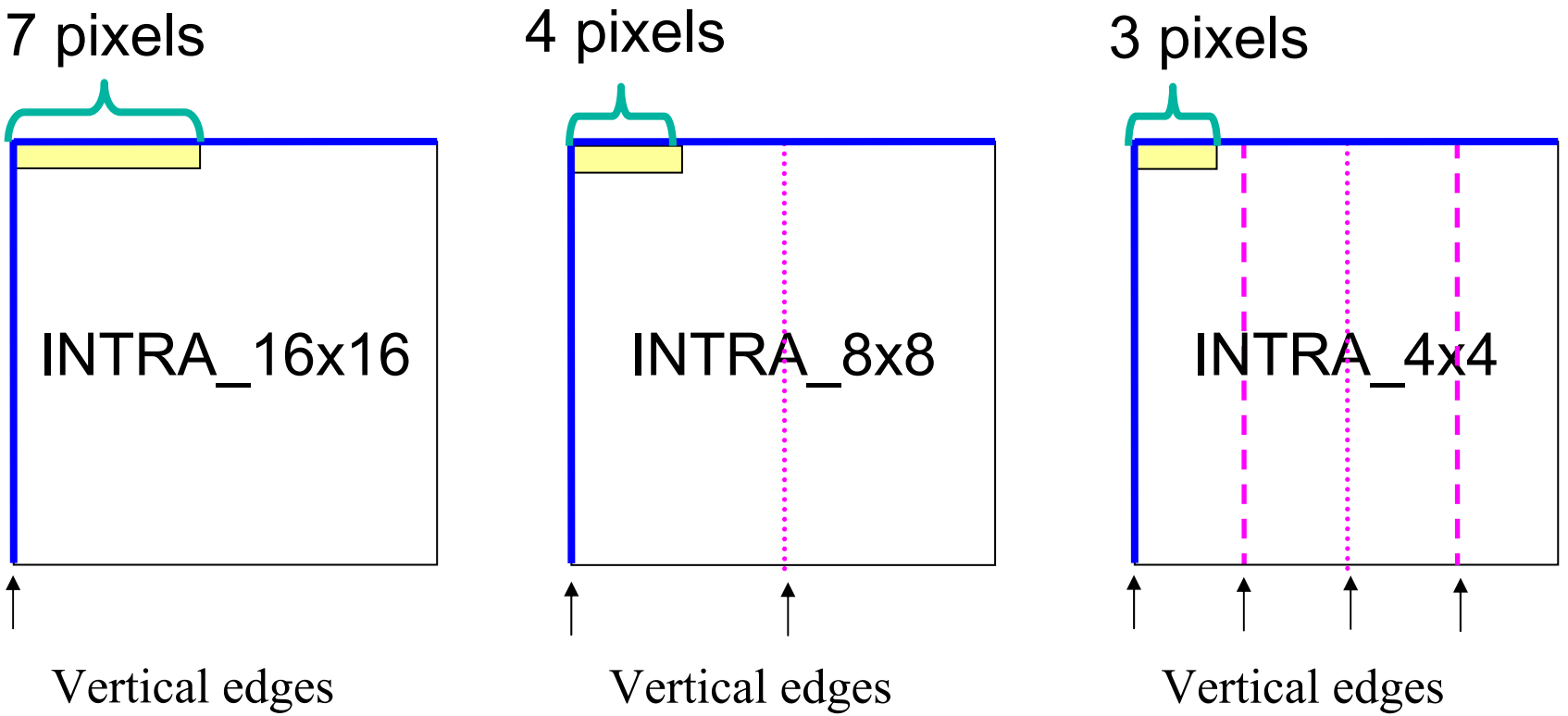
- Significant banding-artifact reduction, while keeping the Rate-Distortion performance

- Recommend a study of the proposed techniques and related technologies

Empowered by Innovation

Appendix: Joint deblocking-debanding filter (1/2)

- # of pixels recursively filtered depends on the associated mb_type and the pixel intensities of edge vicinities.



Appendix: Joint deblocking-debanding filter (2/2)

- Up to 22 pixels are updated by the proposed edge filtering
  - Additional 4 pixels compared to the conventional one; about 22% increase in computational complexity