<1st JCT-VC meeting @Dresden, DE, April 2010>

[A127] Description of video coding technology proposal by ETRI

Hui Yong KIM

hykim5@etri.re.kr

Realistic Media Research Team

17 April 2010 ETRI ***Page total: 24***

Contents

Tool Summary

ETRI Tools

- *Extended block size (32x32) for Intra Slices*
- *Extended intra prediction (32x32, 16x16)*
- *Extended MDDT (32x32, 16x16)*
- *Recursive intra prediction (RIP)*
- *Adaptive filtering process (AFP)*
- *Simplified deblocking filter (SDF)*
- *Enhanced quadtree-based adaptive loop filter (E-QALF)*

KTA2.3 Tools

- *Extended block sizes (64x64, 64x32, 32x64, 32x32, 32x16, 16x32) for Inter Slices*
- *Enhanced adaptive interpolation filter (E-AIF)*
- *Rate-distortion optimized quantization (RDO-Q)*
- *NewOffset for weighted prediction*
- *Mode dependent directional transform (MDDT) (4x4, 8x8, 16x16)*

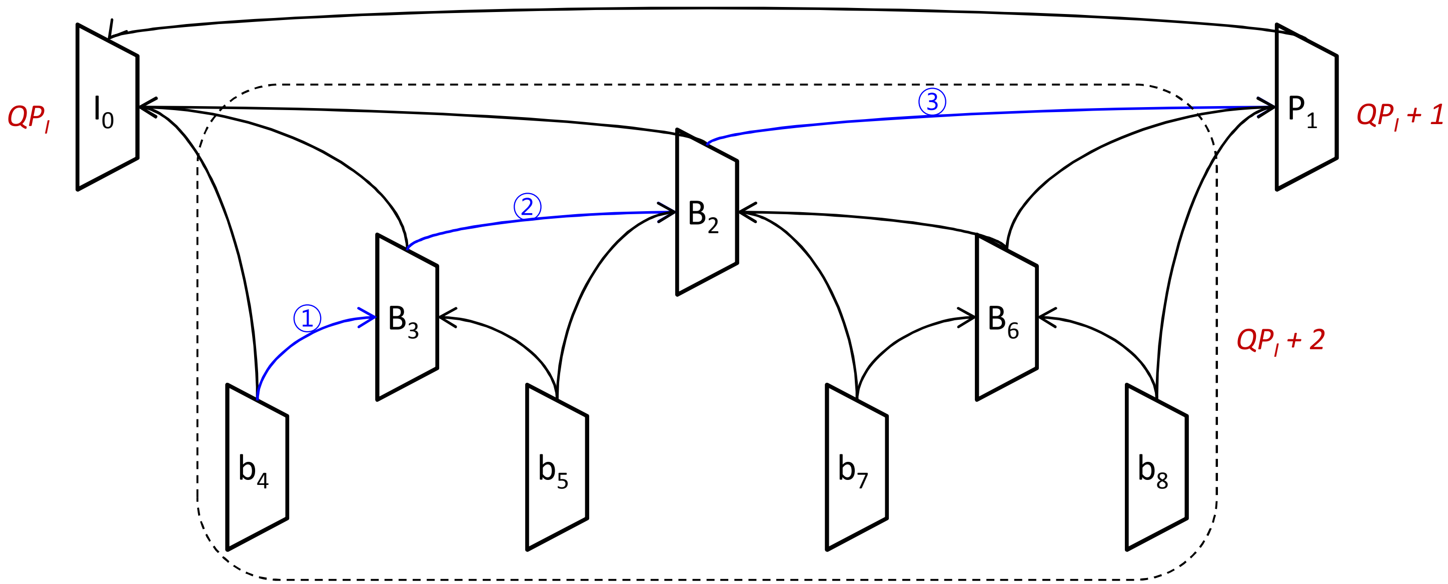
H.264/AVC HP Tools

- *Directional line-based intra prediction (4x4, 8x8, 16x16)*
- *Variable block size motion compensation (4x4 ~ 16x16)*
- *4x4 & 8x8 integer transform, 2x2 & 4x4 Hadamard transform*
- *Scalar quantization*
- *Bi-predictive slice (B-Slice)*
- *In-loop deblocking filter*
- *Weighted prediction*
- *Context adaptive binary arithmetic coding (CABAC)*

Prediction Structure

❑ Constraint Set 1

- ❖ Hierarchical B with 4 levels (similar to Alpha anchor)
  - ➢ open GOP with length=8
  - ➢ max 4 reference frames (B: max 2 per list, P: max 4) not from the lower levels
  - ➢ max 3 frames reordering
- ❖ Structural delay less than 8 pictures
- ❖ Static QPs with possible 1-time change for each seq.
- ❖ Random access intervals less than 1.1 seconds
  - ➢ I-picture every 24/32/48/64 pictures (for 24/30/50/60 fps seq. respectively)

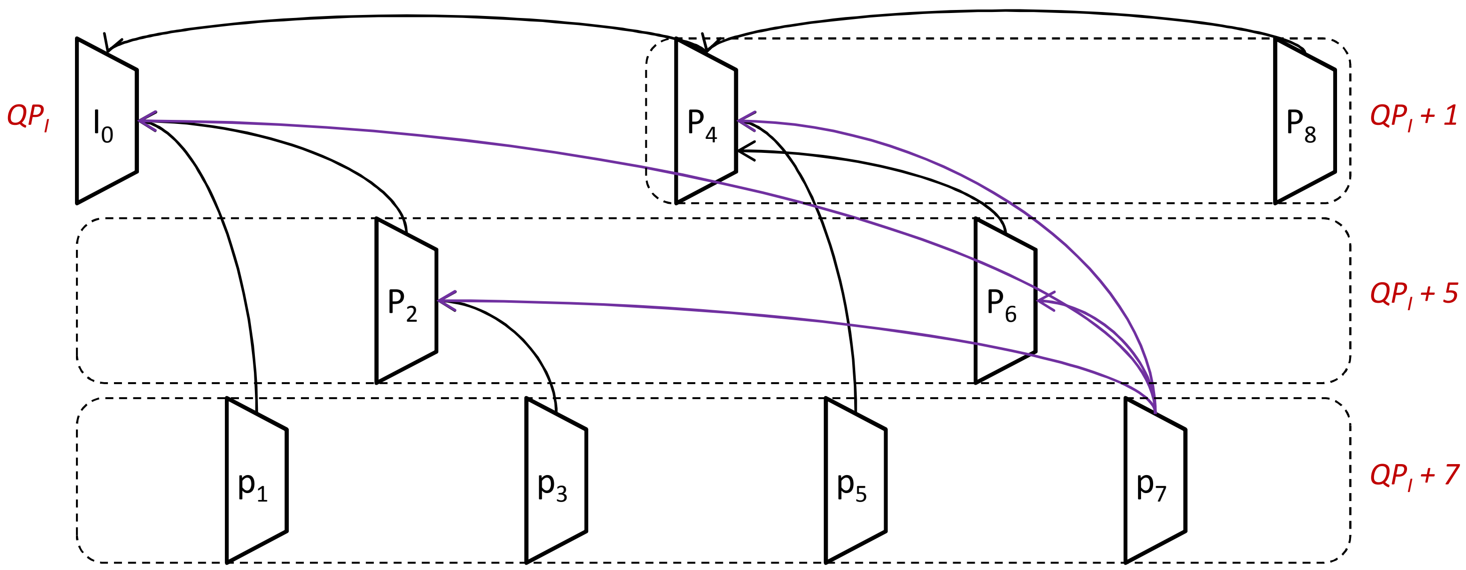


- *Upper case: reference, lower case: non reference*
- *Subscripted numbers: decoding order*

(cont'd)

❑ Constraint Set 2

- ❖ Hierarchical P with 3 levels (similar to Beta anchor)
  - ➢ open GOP with length=4
  - ➢ max 4 reference frames (not from the lower levels)
  - ➢ no picture reordering
- ❖ Static QPs with possible 1-time change for each seq.



- *Upper case: reference, lower case: non reference*
- *Subscripted numbers: decoding order*

Inter Coding

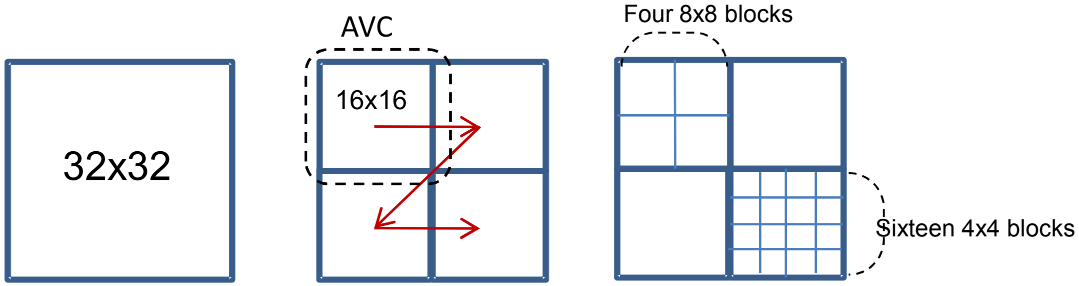
- **Some KTA 2.3 Tools are employed**
  - 64x64 SMB (Super-Macroblock)
    - Ref.) "COM16-C123," Jan. 2009, Qualcomm
  - E-AIF (Enhanced Adaptive Interpolation Filter)
    - Ref.) "VCEG-AJ30," Oct. 2008, Qualcomm
  - Weighted Prediction with NewOffset
    - Ref.) "COM16-C463," Apr. 2008, Qualcomm

- **Others are based on AVC HP (High Profile)**
  - EPZS fast motion estimation (search range: 128x128) from JM11.0
  - etc…

Intra Coding - SMB

❑ 32x32 SMB (Super-Macroblock)

- ❖ Based on "VCEG-AL28", July 2009, ETRI
- ❖ 32x32 SMB structure as an extension of 16x16 MB structure of AVC



- ❖ *I_32x32* prediction
  - ➢ 4 modes similar to AVC *I_16x16* prediction (vertical, horizontal, DC, and plane mode)
  - ➢ Applied to both Luma (32x32) and Chroma (16x16)

- ❖ *I_16x16* Prediction
  - ➢ Luma: 9 modes similar to AVC *I_8x8* (or *I_4x4*) prediction
  - ➢ *mb_type* and *cbp* are coded separately. MPM syntax is adopted for pred_mode.

- ❖ Applied only to I-Slices
  Cf.) In *P/B-Slices*, 64x64 SMB structure of KTA 2.3 is used.

Intra Coding - RIP

❑ Recursive Intra Prediction (RIP)

❖ Recursive predictions along the directional lines

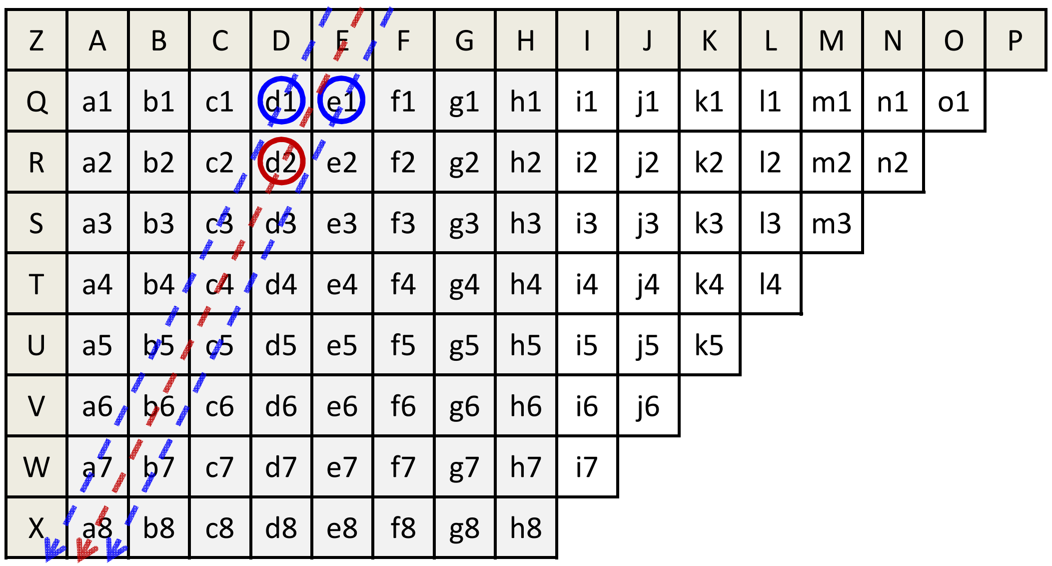
(1) the first pixels are predicted from the reconstructed reference pixels (AVC)

(2) the n^{th} pixels are predicted from the $(n\text{-}1)^{th}$ predicted pixels

❖ Applied only to

- 6 directional modes (excluding vertical, horizontal and DC)
- *I_8x8* and *I_16x16* Luma in *I-Slices*

E.g.) *I_8x8* Vertical-Left



<1st stage> (AVC)

- d1 = (D+E+2)>>2
- e1 = (E+F+2)>>2
- ……

<2nd stage>

- d2 = (d1+e1+2)>>2
- e2 = (e1+f1+2)>>2
- ……

<3rd stage>

- d3 = (d2+e2+2)>>2
- e3 = (e2+f2+2)>>2
- ……

:

<8th stage>

- ……

Intra Coding - AFP

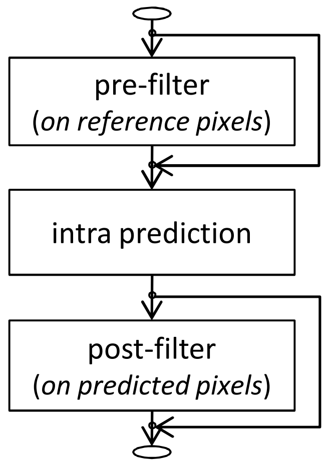
❑ Adaptive Filtering Process (AFP)

(1) Pre-filter before prediction (similar to AVC HP *I_8x8*)

- ➢ 3-tap (¼, ½, ¼) filtering on the reconstructed reference

(2) Post-filter after prediction

- ➢ recursive 3-tap (¼, ½, ¼) filtering on the predicted pixels
- ➢ vertical filtering followed by horizontal filtering



- ❖ RDO-based filter on/off decision
- ❖ Applied only to *I_8x8* and *I_16x16* Luma in *I-Slices*

E.g.) *I_8x8* Post-filter

Z	A	B	C	D	E	F	G	H
Q	a1	b1	c1	d1	e1	f1	g1	h1
R	a2	b2	c2	d2	e2	f2	g2	h2
S	a3	b3	c3	d3	e3	f3	g3	h3
T	a4	b4	c4	d4	e4	f4	g4	h4
U	a5	b5	c5	d5	e5	f5	g5	h5
V	a6	b6	c6	d6	e6	f6	g6	h6
W	a7	b7	c7	d7	e7	f7	g7	h7
X	a8	b8	c8	d8	e8	f8	g8	h8

(→ H: horizontal; ↓ V: vertical)

<Vertical>

$V[a1] = (A + 2\cdot a1 + a2 + 2) \gg 2$

$V[a2] = (V[a1] + 2\cdot a2 + a3 + 2) \gg 2$

. . .

$V[a8] = (V[a7] + 3\cdot a8 + 2) \gg 2$

$V[b1] = (B + 2\cdot b1 + b2 + 2) \gg 2$

$V[b2] = (V[b1] + 2\cdot b2 + b3 + 2) \gg 2$

:

:

<Horizontal>

$H[a1] = (Q + 2\cdot V[a1] + V[b1] + 2) \gg 2$

$H[b1] = (H[a1] + 2\cdot V[b1] + V[c1] + 2) \gg 2$

. . .

$H[h1] = (H[g1] + 3\cdot V[h1] + 2) \gg 2$

$H[a2] = (R + 2\cdot V[a2] + V[b2] + 2) \gg 2$

:

:

<AFP modes>

MB type	AFP modes	Pre-filter	Post-filter	
I_8x8	1	*on*	*on*	
	2	*on*	*off*	: AVC HP
	3	*off*	*on*	
	4	*off*	*off*	
I_16x16	1	*off*	*on*	
	2	*off*	*off*	

Transforms & Quantization

❑ Transforms in I-Slices

- ❖ Based on KTA 2.3 MDDT(“VCEG-AJ24”, Oct. 2008, Qualcomm)
- ❖ New MDDT kernels for I_32x32(4-modes) & I_16x16(9-modes)

❑ Transforms in P/B-Slices

- ❖ KTA 2.3 is used without modification.

Slice type	Luma		Chroma	
	Block size	Transform	Block size	Transform
I	32x32	MDDT: 32x32	16x16	8x8 ICT + 2x2 HT
	16x16	MDDT: 16x16	8x8	4x4 ICT + 2x2 HT
	8x8	MDDT: 8x8		
	4x4	MDDT: 4x4		
P or **B**	64x64, 64x32, 32x64, 32x32, 32x16, 16x32	ICT: 16x16, 8x8, 4x4	8x8	4x4 ICT + 2x2 HT
	16x16	ICT: 16x16, 8x8, 4x4 (MDDT: 16x16 for I)		
	16x8	ICT: 16x8, 8x8, 4x4		
	8x16	ICT: 8x16, 8x8, 4x4		
	8x8	ICT: 8x8, 4x4 (MDDT: 8x8 for I)		
	8x4, 4x8, 4x4	ICT: 4x4 (MDDT: 4x4 for I)		

*MDDT: Mode-Dependent Directional Transform
*ICT: Integer Cosine Transform
*HT: Hadamard Transform

❑ Quantization

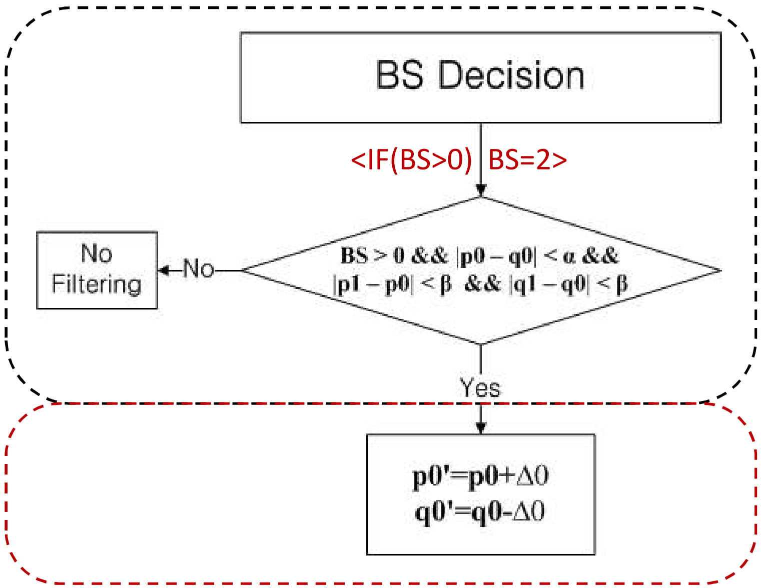
- ❖ RDO-Q from KTA 2.3 is employed for all the transform types.
- ❖ Ref.) “VCEG-AI36,” Jul. 2008, Qualcomm

In-Loop Filters

❑ Simplified Deblocking Filter (SDF)

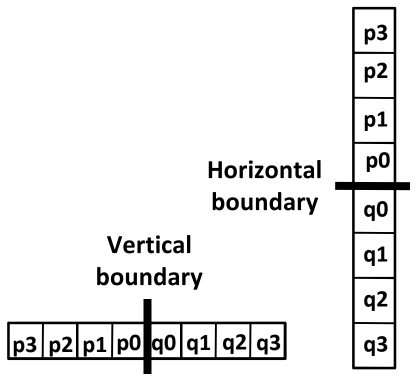
- ❖ Simplified version of AVC deblocking filter
  - ➢ only the boundary pixels are filtered (p0 & q0)
  - ➢ reduced # of BS types: if BS>0, then BS is set to 2
- ❖ Ref.) "VCEG-AJ17" Oct. 2008, Kwangwoon Univ.

<on/off decision from AVC>



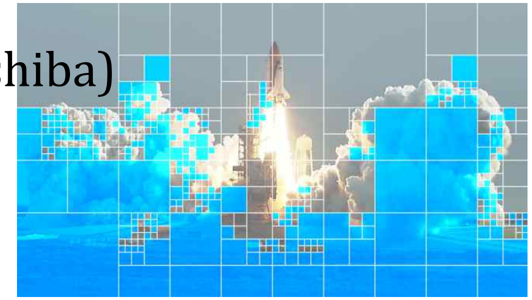
<simplified filtering>

$$\Delta 0 = \mathrm{Min}(\mathrm{Max}(-c0, \Delta 0i), c0)$$
$$\Delta 0i = ((q0-p0) \ll 2 + (p1-q1) + 4) \gg 3$$
$$c0 = c1 + (|p2-p0| < \beta) + (|q2-q0| < \beta)$$
$$c1 = \mathrm{ClipTable}[QP, BS]$$

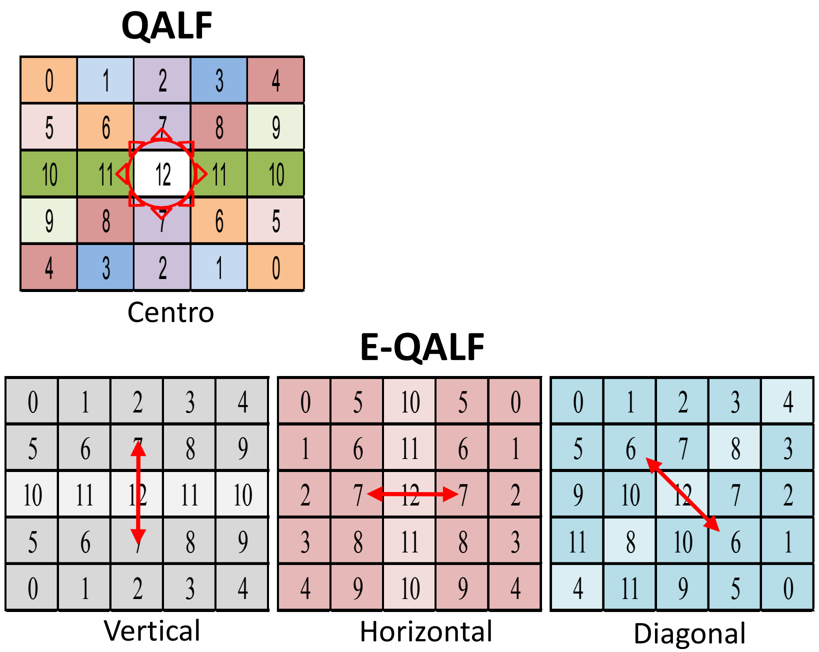


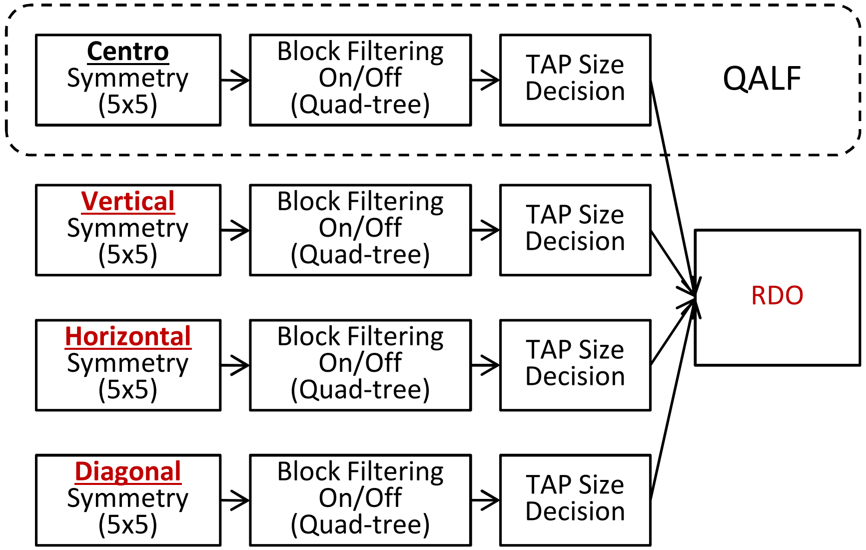
(cont'd)

❑ Enhanced QALF (E-QALF)



- ❖ QALF from KTA 2.3 ("VCEG-AK22", Apr. 2009, Toshiba)
  - ➢ loop filter based on Wiener filter
  - ➢ quadtree-based filter on/off
- ❖ E-QALF adds 3-more symmetries to QALF
  - ➢ Competition based on RDO at the last stage
  - ➢ improved RD performance without increasing decoder complexity





Entropy Coding

- **CABAC in AVC HP is employed.**
  - For the newly added Syntax Elements (SEs), the binarization procedure and the context models were designed similarly to AVC.

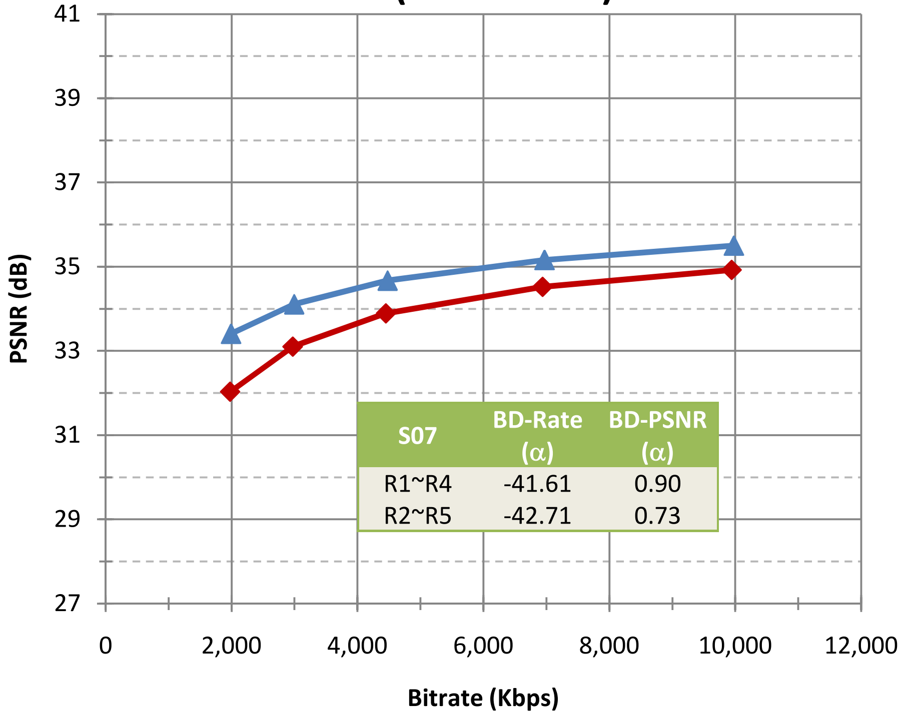
RD Performance

❑ Constraint Set 1 (relative to Alpha)

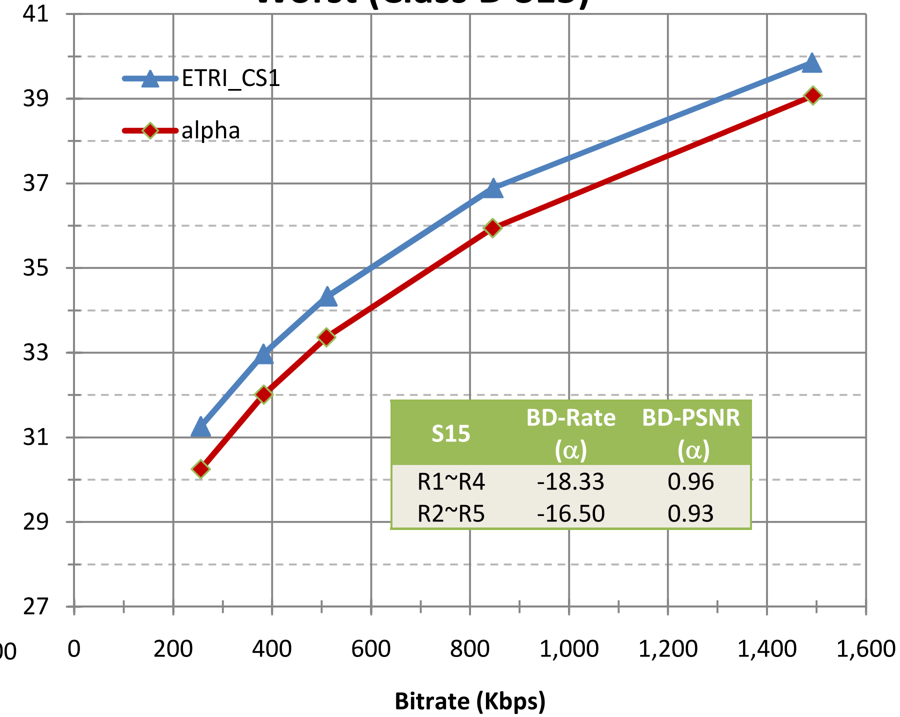
Class	BD-Rate (α)	BD-PSNR (α)
A	-24.89	1.17
B	-33.74	1.11
C	-28.68	1.36
D	-23.45	1.12
Avg.	-28.47	1.19

(Only Luma results are shown here.)

Best (Class B S07)



Worst (Class D S15)

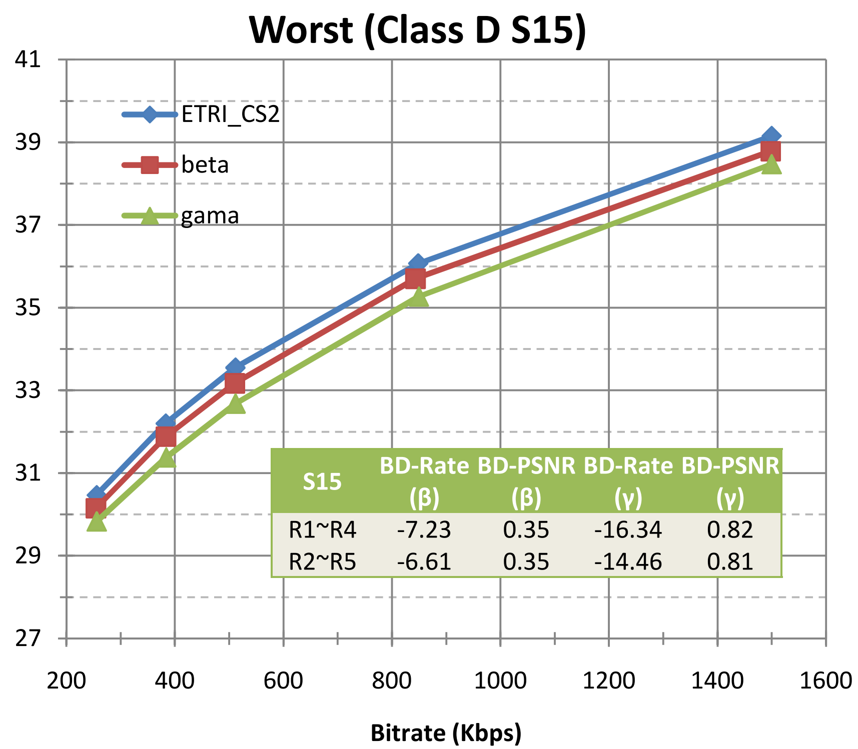


(cont'd)

❑ Constraint Set 2 (relative to Beta & Gamma)

Class	BD-Rate (β)	BD-PSNR (β)	BD-Rate(γ)	BD-PSNR (γ)
B	-30.58	0.77	-48.94	2.02
C	-20.36	0.89	-38.72	1.90
D	-10.85	0.47	-33.23	1.62
E	-34.96	1.60	-51.80	2.78
Avg.	-25.51	0.94	-45.86	2.17

(Only Luma results are shown here.)



Subjective Quality Comparison

- C1 - Class A S01 & S02 @R1(2.5Mbps)

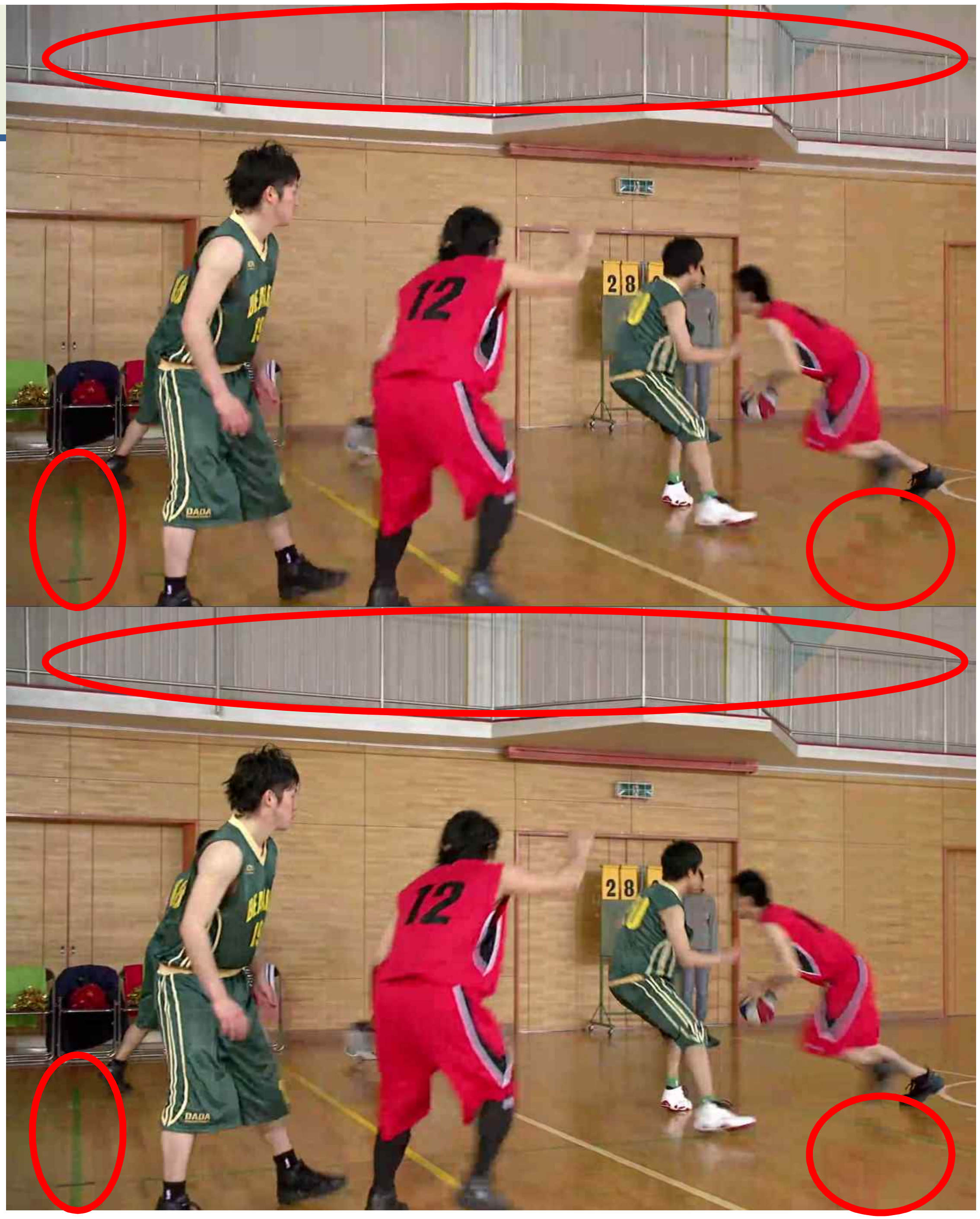
(cont'd)

- C1 – Class C S08 @R1(384Kbps)

(co

- C2 - Class B S03 @R1(1Mbps)

❑ C2 - Class B S06 @R1(2Mbps)



Complexity

❑ Time Measurement

- ❖ '*ntimer*' for Windows systems were used
- ❖ All the measurements were performed on the same platform
  - ➢ Microsoft Windows Vista Business K 64bit
  - ➢ Intel Xeon CPU X5482 (3.2GHz 2-CPUs, 8 Core) with 8GB RAM
- ❖ Both anchors and proposed are compiled into 64-bit binary executables

❑ Encoding Time

Constraint Set 1

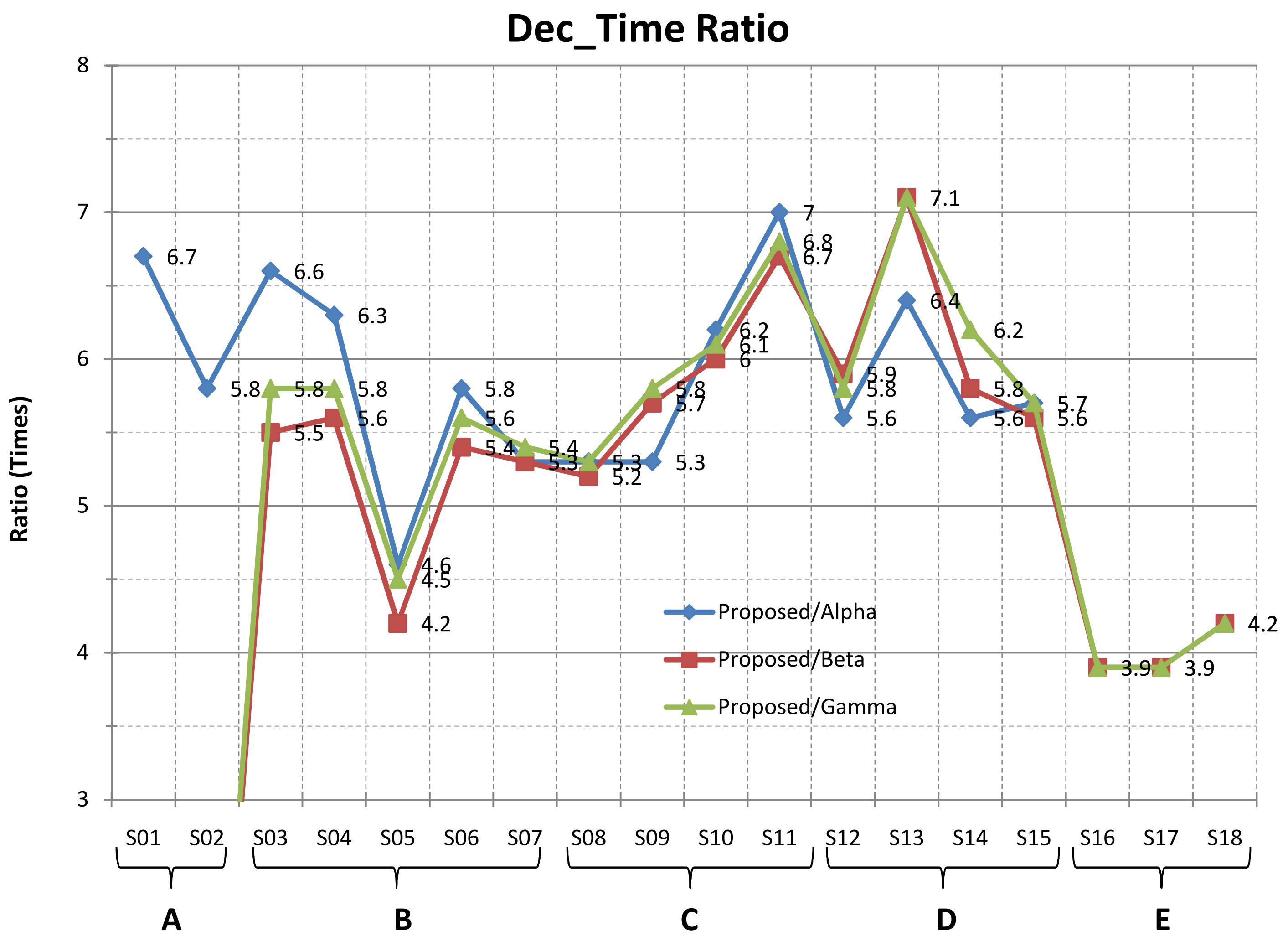
Class	Seq.	Ave.	fps	EncTime/DecTime
A	S01	113,256	0.0013	427
	S02	132,799	0.0011	578
B	S03	117,636	0.0020	487
	S04	107,421	0.0022	484
	S05	242,098	0.0021	736
	S06	245,059	0.0020	548
	S07	265,227	0.0023	542
C	S08	49,832	0.0100	944
	S09	67,189	0.0089	985
	S10	53,531	0.0093	817
	S11	34,277	0.0088	797
D	S12	16,761	0.0298	1133
	S13	15,854	0.0378	734
	S14	16,633	0.0301	1139
	S15	10,937	0.0274	968

Constraint Set 2

Class	Seq.	Ave. (sec)	fps	EncTime/DecTime
B	S03	94,373	0.0025	510
	S04	88,644	0.0027	492
	S05	195,300	0.0026	703
	S06	194,591	0.0026	531
	S07	203,439	0.0029	491
C	S08	40,305	0.0124	841
	S09	53,004	0.0113	822
	S10	41,442	0.0121	735
	S11	32,702	0.0092	884
D	S12	13,621	0.0367	1032
	S13	11,073	0.0542	589
	S14	12,780	0.0391	968
	S15	9,917	0.0302	992
E	S16	81,678	0.0073	716
	S17	81,773	0.0073	711
	S18	81,981	0.0073	669

(cont'd)

❑ Decoding Time (relative to Anchors)



Software Implementation

- Based on JM11.0KTA2.3.

- **Program Language**
  - C

- **Compilation**
  - Platform: Microsoft Windows 7 x64 edition
  - Compiler: Microsoft Visual Studio 2008
    - both 64- and 32-bit compiled versions are submitted

- **Low-level programming optimization**
  - No.

- **Usage of external video libraries**
  - No.

Summary

- **ETRI's CODEC is based on the KTA 2.3**
  - 64x64 SMB for Inter-Slices
  - MDDT
  - E-AIF
  - RDO-Q
  - NewOffset

- **Proposed tools are:**
  - 32x32 SMB for I-Slices
  - 4-modes I_32x32 & 9-modes I_16x16 prediction
  - 32x32 and 16x16 MDDT kernels
  - Recursive Intra prediction for the 6 directional modes (RIP)
  - Adaptive pre-/post-filtering for Intra prediction (AFP)
  - Simplified deblocking filter (SDF)
  - Enhanced QALF (E-QALF)

- **RD Performance**
  - Under CS1 (against Alpha)
    - min -23.45% (Class D) and max -33.74% (Class B) average BD-Rate are achieved.
  - Under CS2
    - min -10.85% (Class D) and max -34.96% (Class E) average BD-Rate are achieved.
    - min - 33.23% (Class D) and max -51.80% (Class E) average BD-Rate are achieved.

Thank You Very Much !

www.etri.re.kr