LG’s proposal for HVC

JCTVC-A110

ByeongMoon Jeon,
SeungWook Park,
JungSun Kim,
JoonYoung Park

LG Electronics

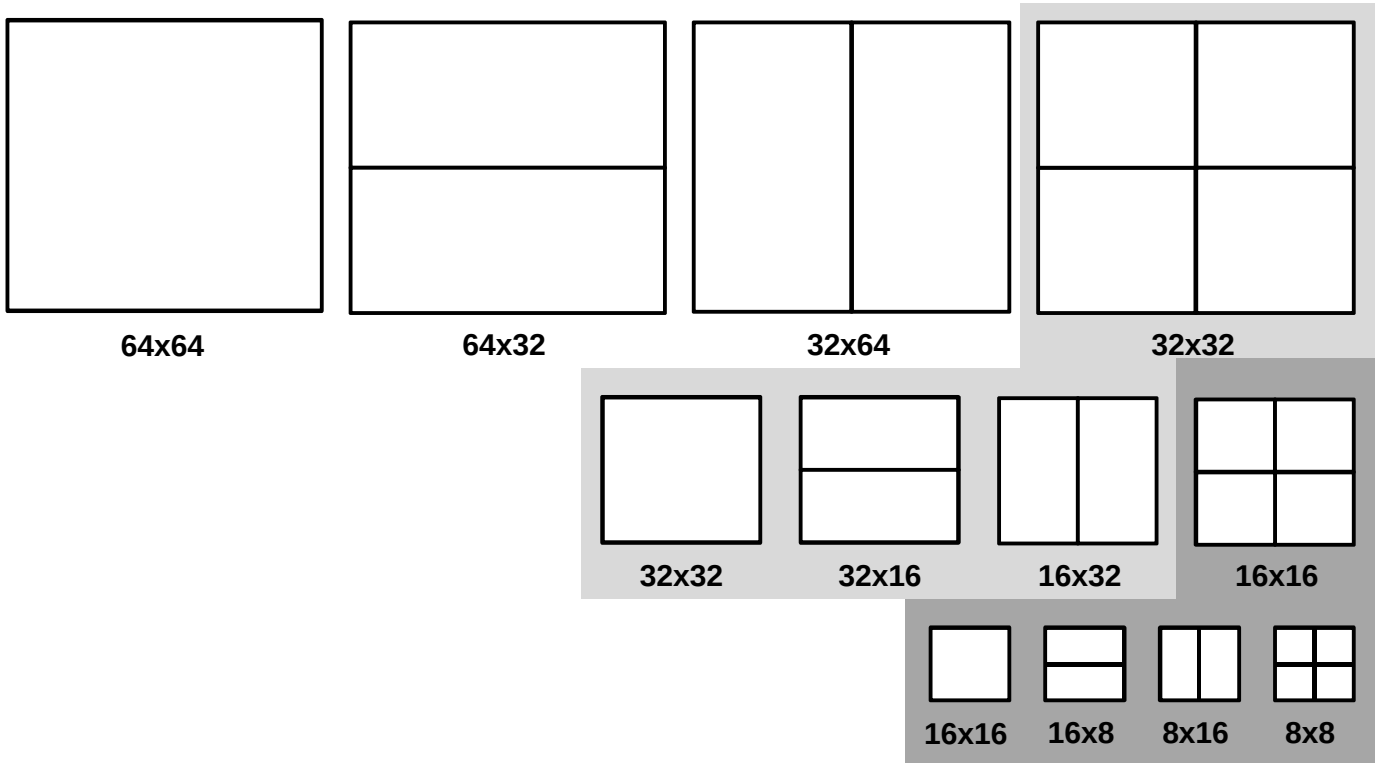
Introduction

- Based on AVC/H.264, the following new features are added
  - **Large macroblock structure (64x64 to 8x8 coding blocks)**
  - **IIMM (Inter-Intra Mixed Mode)**
  - **Skip mode with variable block size**
  - **SPMV (Scaled Predicted Motion Vector)**
  - **Template-based IC(Illumination Compensation)**
  - **Modified MVC (Motion Vector Competition)**
  - **SIFO of KTA**
  - **New intra prediction types (I_32x32, I_mixed, ..,etc)**
  - **New chroma intra prediction mode (chroma estimation mode)**
  - **Border handling scheme**
  - **ADF (Adaptive Deblocking Filter)**
  - **QALF of KTA**
  - **MDDT with modified kernels**
  - **Adaptive scan ordering**
  - **AWR (Adaptive Warped Reference)**
  - **PAIF (Parametric Adaptive Interpolation Filter)**
  - **MVC with B skip/direct**
- Proposed model has Substantially increased compression capability relative to AVC/H.264

Macroblock Structure

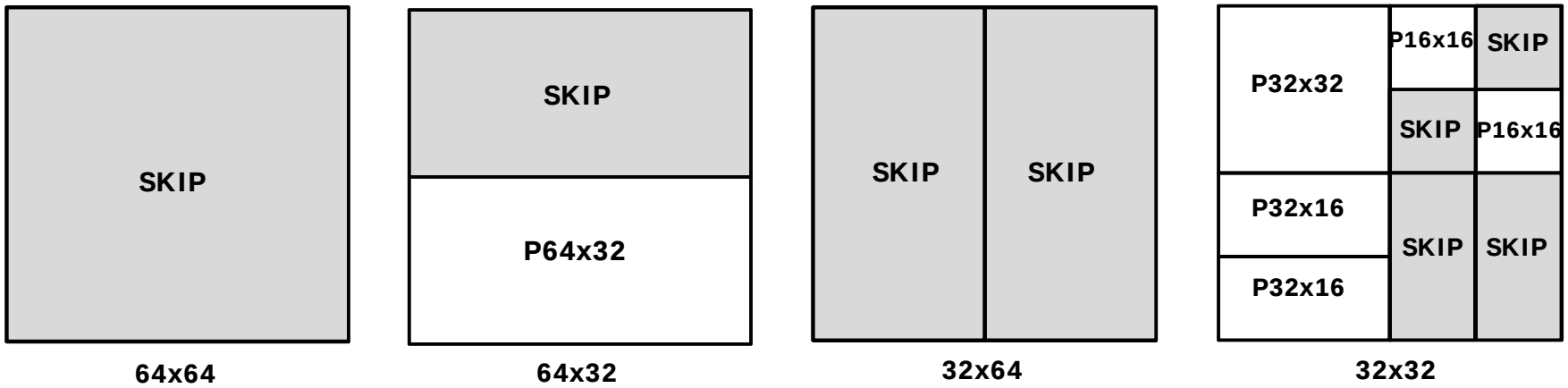
v Macroblock structure for inter prediction

- § Macroblock unit size is 32x32
- § partitioned into several shapes of motion-compensated blocks
- § Can be clustered up to 64x64



Macroblock Structure (Cont'd)

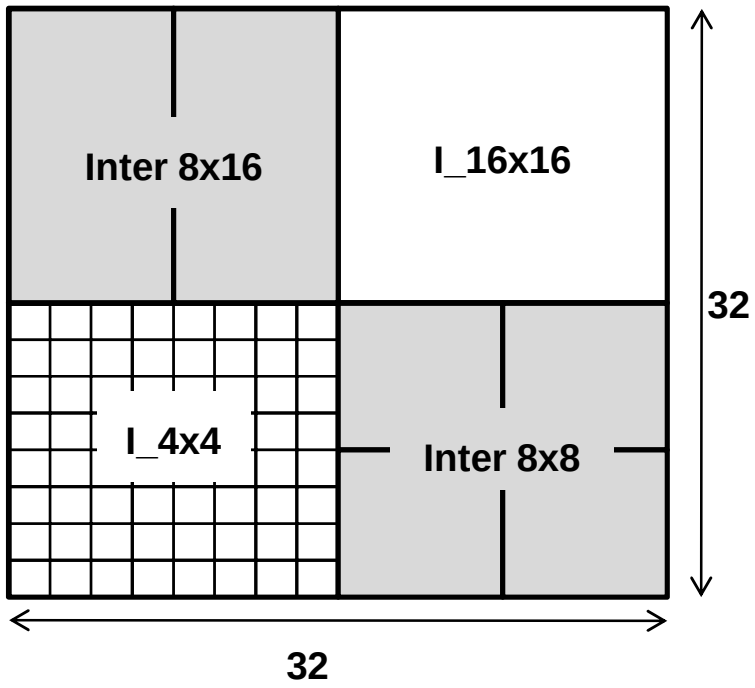
- **Skip mode with variable block size**
  - Applies to the following blocks partitions
    - 64x64, 64x32, 32x64, 32x32 → efficient for high resolution
    - 32x16, 16x32, 16x16 → efficient for low resolution
  - Increases the portion of skip mode
    - Increases coding efficiency



Macroblock Structure (Cont'd)

v IIMM (Inter-Intra Mixed Mode)

- § A macroblock (32x32) has both inter and intra submacroblocks
- § 1 bit flag in each submacroblock (16x16) to signal inter or intra
- § Same structure with AVC/H.264 **à** efficient for low resolution

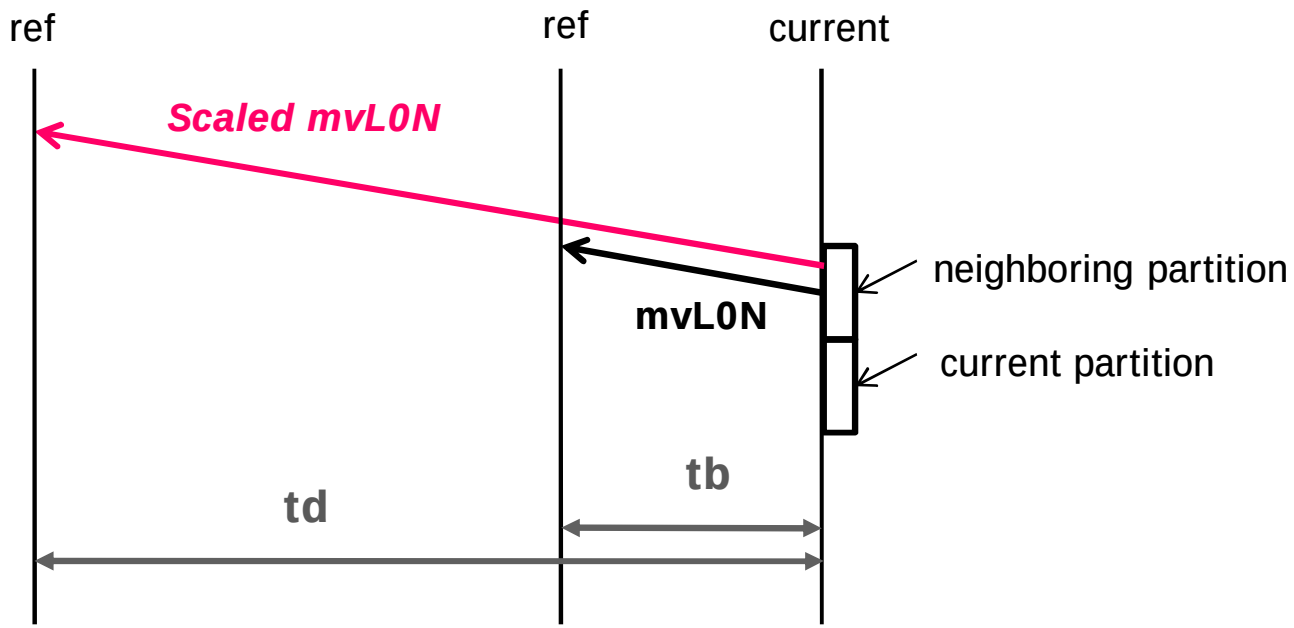


Inter prediction

v SMVP (Scaled motion vector predictor)

§ The motion vectors of the neighboring blocks is scaled according to the temporal distance between current and reference pictures

Scaled mvLXN = (td / tb) * mvLXN



Inter prediction (Cont'd)

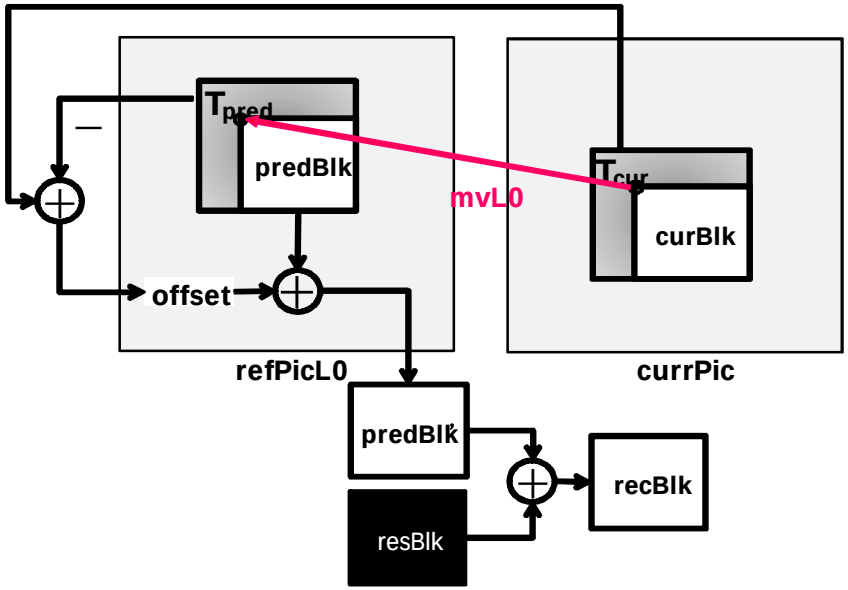
v Template-based IC (Illumination Compensation)

§ To compensate the illumination change between pictures

§ Offset value representing the illumination change between current and predicted block is derived at decoder

- Ø **Offset is set to the difference value between the DC values of T_{cur} and T_{pred}**

§ Offset value is added to the predicted block to compensate illumination change



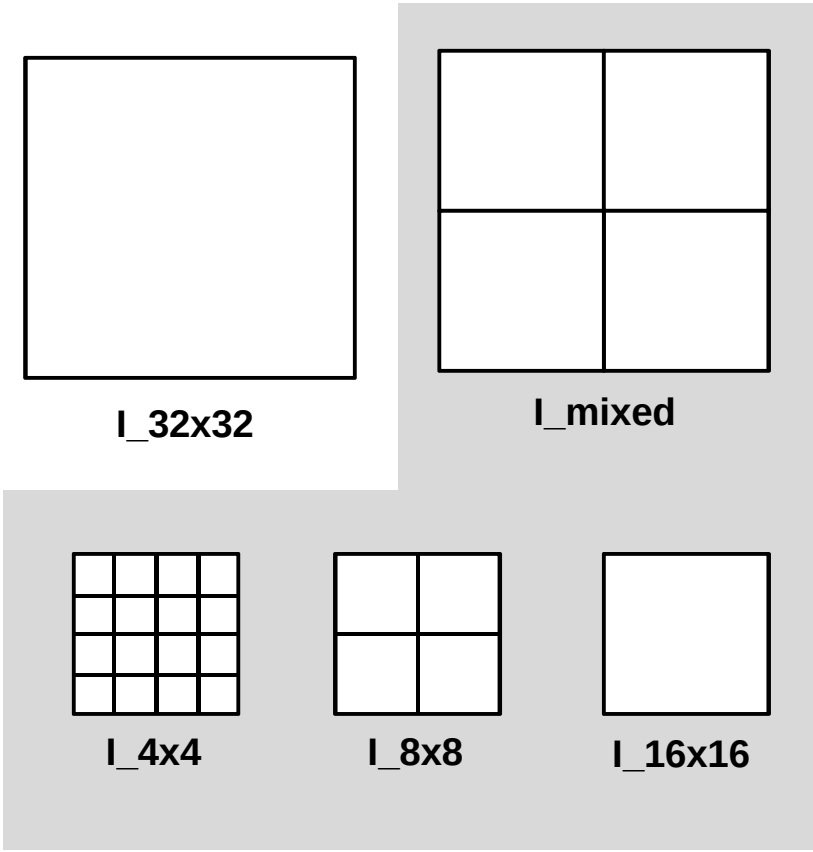
Inter prediction (Cont'd)

- **Modified MVC (Motion Vector Competition)**
  - MVC of KTA is modified to be compatible with the new tools
    - Combined with SPMV(Scaled predicted motion vector)

- **SIFO(Switched Interpolation Filter Offset)**
  - The single pass SIFO of KTA is employed

Intra-frame prediction

- **Intra prediction types**
  - A macroblock has one of two intra macroblock modes
    - **I_32x32, I_mixed**
  - Each submacroblock has one of three submacroblock modes
    - **I_4x4, I_8x8, I_16x16**



Intra-frame prediction (cont'd)

v I_32x32

- Four prediction modes (vertical, horizontal, DC, plane) of AVC/H.264 is used for 32x32 block

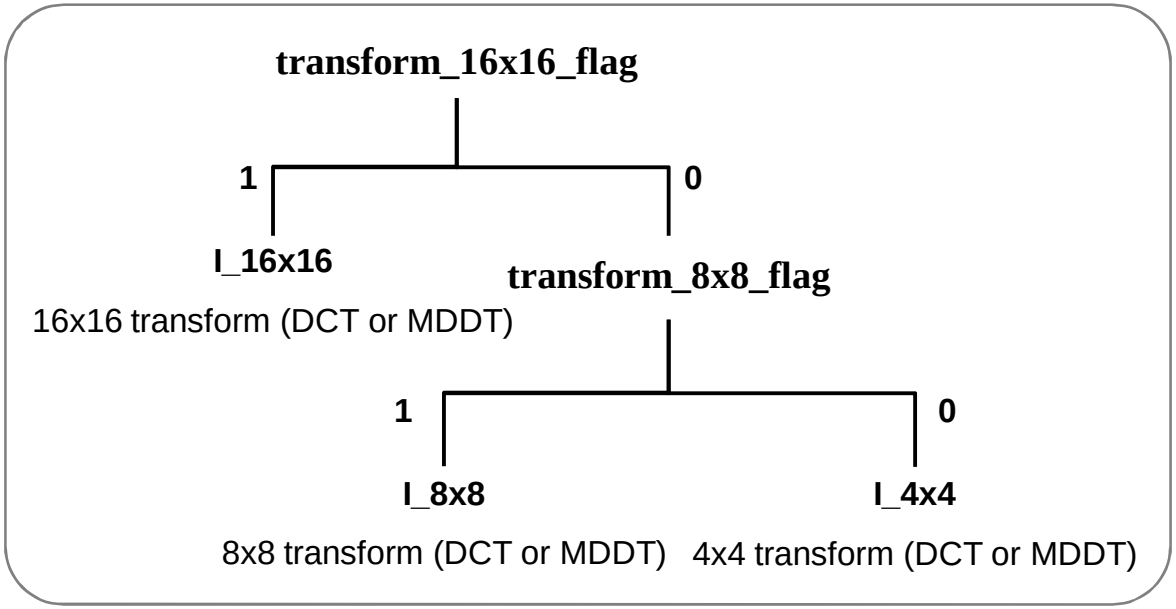
Intra32x32PredMode	Name of Intra32x32PredMode
0	Intra_32x32_Vertical
1	Intra_32x32_Horizontal
2	Intra_32x32_DC
3	Intra_32x32_Plane

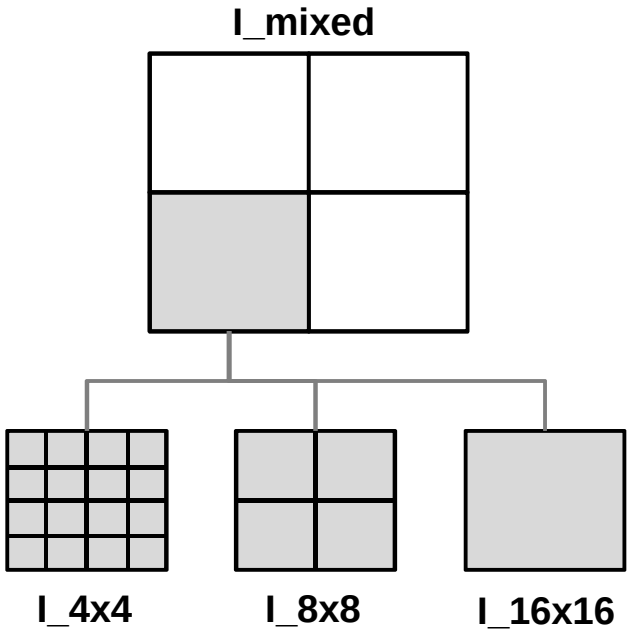
Intra-frame prediction (Cont'd)

v I_mixed

- § Each submacroblock can be coded one of three prediction modes,
  - Ø I_4x4, I_8x8, I_16x16
  - Ø Each mode has one of 9 prediction mode of AVC/H.264

- § Two flags (transform_16x16_flag , transform_8x8_flag) signal prediction type and transform size of a submacroblock





Intra-frame prediction (Cont'd)

v Chroma intra prediction

- § Prediction unit size is 8x8
- § New chroma prediction mode (chorma estimation mode) is added to AVC/H.264 modes (DC, vertical, horizontal, plane)
- § Predicted chroma samples are derived from the reconstructed luma samples based on linear model

$$\mathbf{pred_C[\ x, y\ ] = \alpha * P_L[\ 2*x, 2*y\ ] + \beta}$$

$$a = \frac{R(P'_L, P'_C)}{R(P'_L, P'_L)}$$

$$b = mean(P'_C) - a * mean(P'_L)$$

$pred_C[\ x, y\ ]$: chroma pixels to be predicted

$P_L[\ 2*x, 2*y\ ]$: subsampled luma pixels

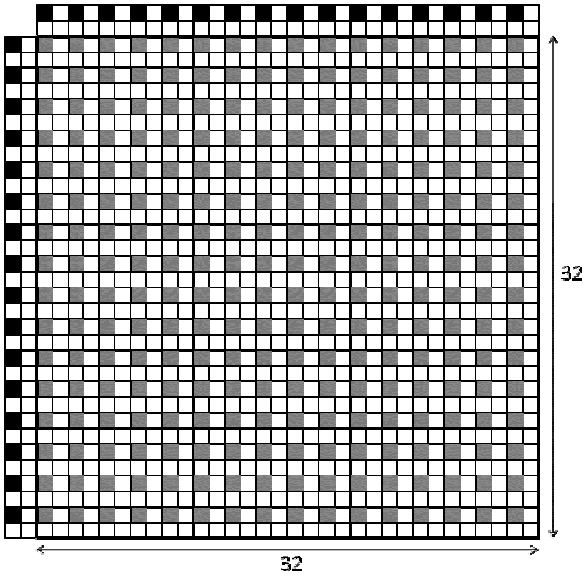
→ gray pixels in luma block

P'_L : pixels subsampled from neighboring pixels of luma block

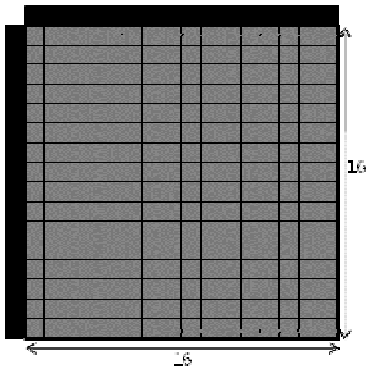
→ black pixels in luma block

P'_C : neighboring pixels of chroma block

→ black pixels in chroma block



Luma block



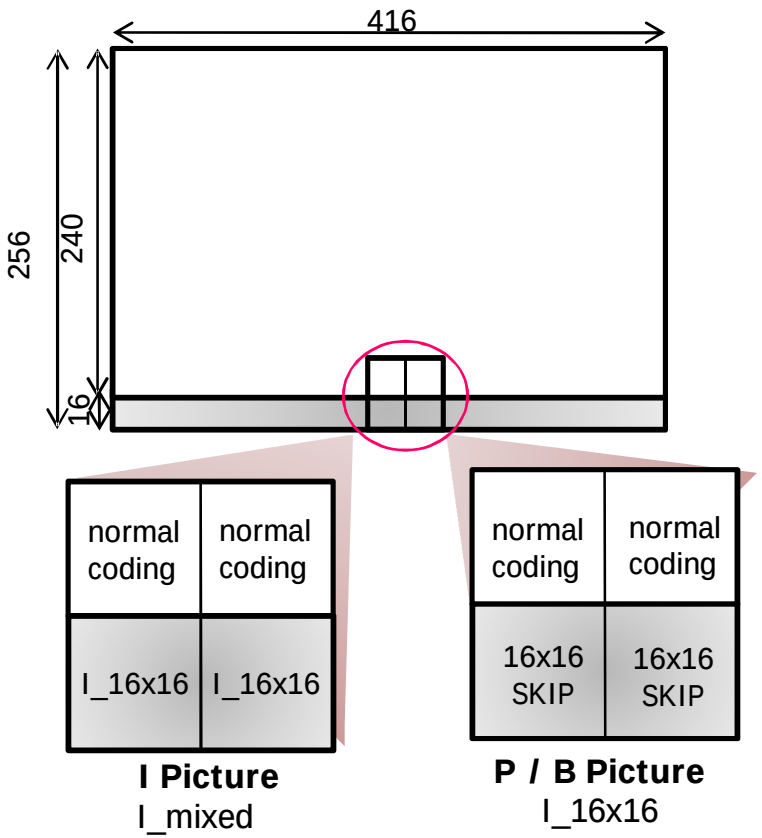
Chroma block

Border handling scheme

v PROPOSED METHOD

- § When a part of macroblock (32x32) lies within the padded region, the macroblock is partitioned into four 16x16 submacroblocks

- § Submacroblocks within the padded region are derived at decoder without any overhead
  - **v P / B Picture**
    - **Ø** 16x16 SKIP
  - **v I Picture**
    - **Ø** I_16x16
    - **Ø** prediction mode depends on its position
      - right border : horizontal mode
      - bottom border : vertical mode



Transform / Scan Order

v Transforms

- § DCT : 4x4, 8x8, 16x8, 8x16, 16x16
- § MDDT : New MDDT kernels

	Luma		Chroma
	Block size	Transform	Transform
Inter	64x64, 64x32, 32x64, 32x32, 32x16, 16x32, 16x16	§ 4x4, 8x8, 16x16	§ 4x4 § 2x2 or 4x4 DC Hadamard depending on skip status
	16x8	§ 4x4, 8x8, 16x8	
	8x16	§ 4x4, 8x8, 8x16	
	8x8	§ 4x4, 8x8	
Intra	I_32x32	§ DC pred mode : 16x16 DCT § Other modes : 16x16 MDDT	§ 4x4 § 4x4 DC Hadamard
	I_mixed (I_16x16, I_8x8, I_4x4)	Same transform size as block size	

v Scan order

- § Adaptive scan order with new initial values

In-loop filters / Entropy Coding

- **In-loop Filters**
  - Adaptive Deblocking Filter (ADF)
    - Adaptive deblocking filter based on ‘Wiener filtering’ scheme
    - Filter coefficients are calculated by resolving MMSE between the original and the reconstructed pictures.
    - ADF or AVC/H.264 deblocking filter is applied adaptively in each slice
  - Quad-tree based Adaptive Loop Filter (QALF)
    - QALF of KTA is employed
- **Entropy Coding**
  - AVC/H.264 CABAC is employed

Other tools

- Tools being developed but not included in the submitted model
- By adopting these tools, further improvements can be expected

AWR (Adaptive Warped Reference)

- Generates a new warped reference picture considering complex motions such as scaling, rotation, sheering, and so on.
- Modeling of a parametric warping function using KLT(Kanade-Lucas-Tomasi) feature tracker
- Refer to JCTVC-021

PAIF (Parametric Adaptive Interpolation Filter)

- Transmits just 5 parameters instead of individual filter coefficients to represent interpolation filter
- Less bits for representing filters and closer to optimal filter than existing AIFs
- Refer to JCTVC-021

MVC (Motion Vector Competition) with B skip/direct

- MVC with B skip/direct is utilized to maximize benefits of MVC scheme

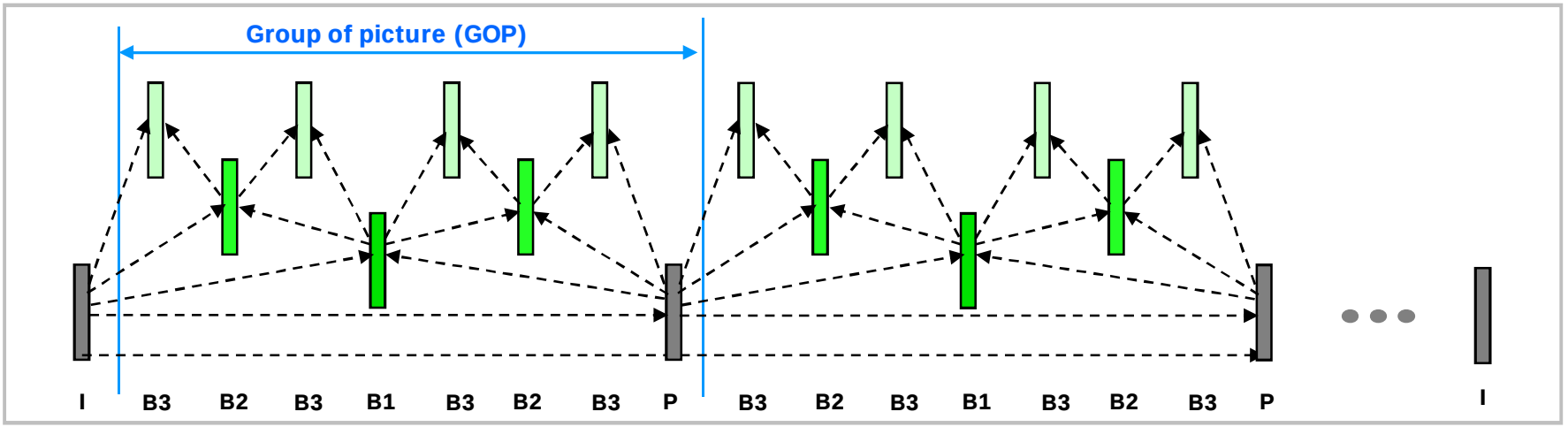
Chroma estimation mode with phase shift

- The new chroma prediction mode (chroma estimation mode) is modified considering phase difference between luma and chroma samples

Coding structure

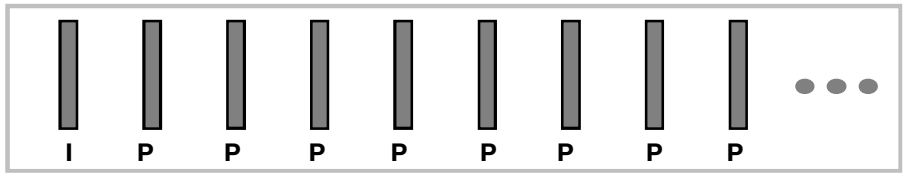
Constraint set 1

- Hierarchical B coding / GOP8/ IDR picture every 1 sec
- Same configuration of alpha anchor except the following changes
  - Number of reference frames for P pictures=5
  - Number of reference frames for B pictures=4 (2 reference pictures allowed for each list)
  - Weighted prediction disabled
  - 16x16, 16x8, 8x16, 8x8, 4x4 transforms enabled



Constraint set 2

- I-P-P-P coding / IDR only at first picture
- Same configuration of gamma anchor except the following changes
  - Number of reference frames for P pictures=5
  - CABAC enabled
  - 16x16, 16x8, 8x16, 8x8, 4x4 transforms enabled



RD performance

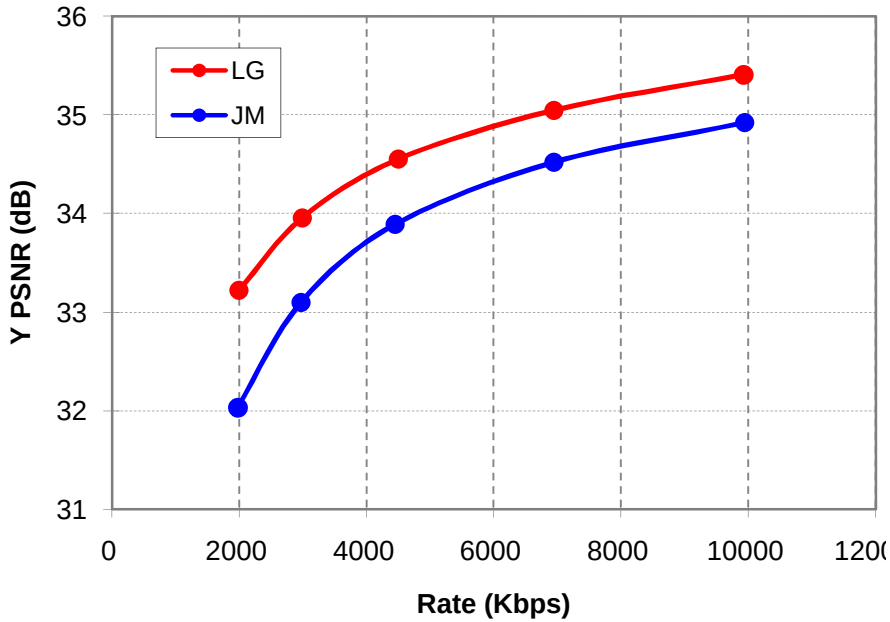
v Constraint Set 1 (alpha)

Class	BD psnr	BD rate
A	1.11	-23.94
B	1.01	-30.57
C	1.26	-26.85
D	0.92	-19.85
Avg.	**1.06**	**-25.83**

Max

BD-PSNR	BD-Rate
0.71	**-36.15**

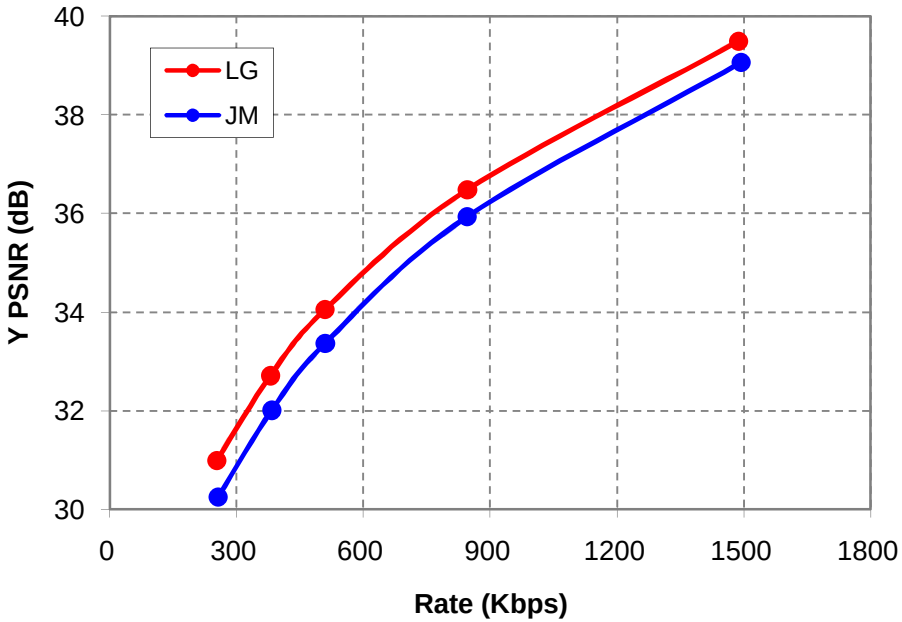
Class B-BQTerrace



Min

BD-PSNR	BD-Rate
0.62	**-11.76**

Class D-RaceHorses



RD performance

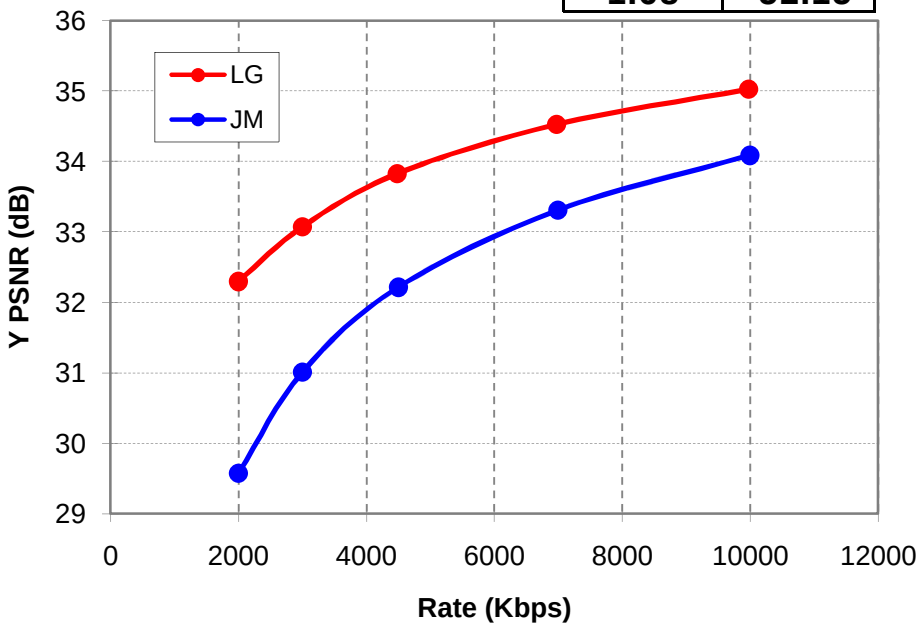
v Constraint Set 2 (gamma)

Class	BD psnr	BD rate
B	1.85	-44.42
C	1.60	-33.65
D	1.15	-25.11
E	2.31	-45.01
Avg.	**1.70**	**-37.01**

Max

Class B-BQTerrace

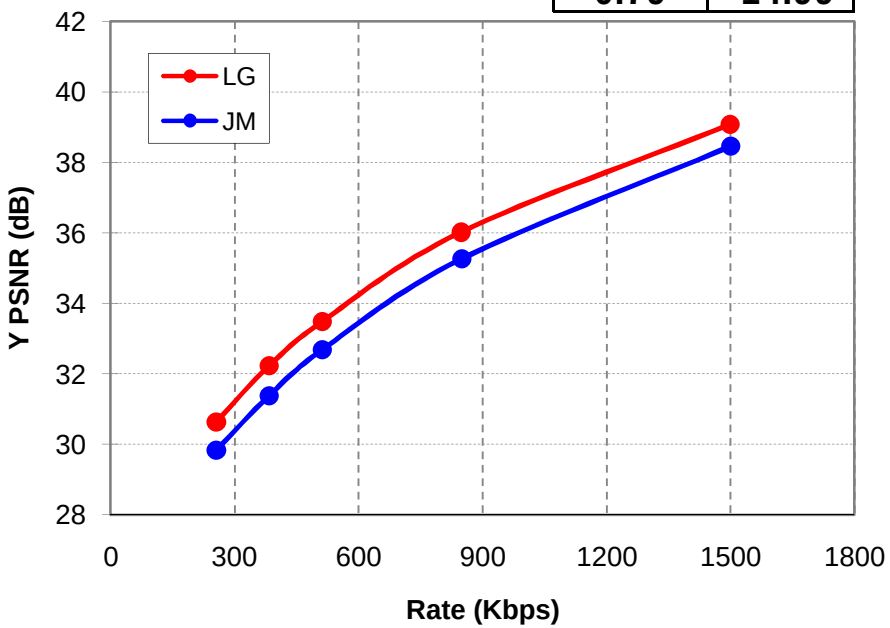
BD-PSNR	BD-Rate
1.68	**-52.15**



Min

Class D-RaceHorses

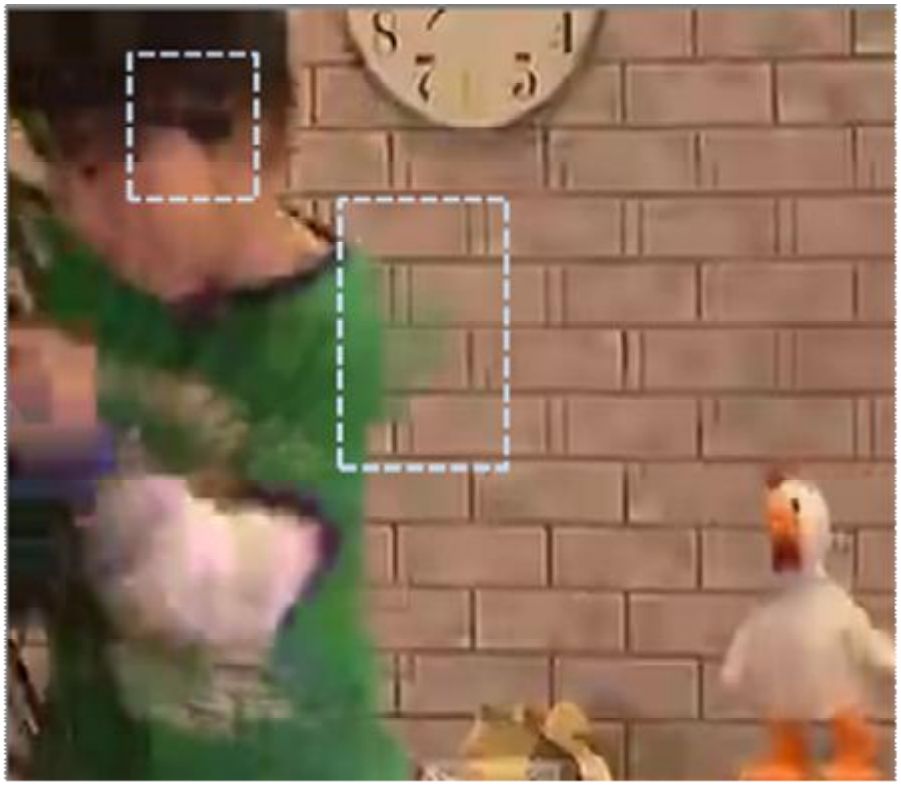
BD-PSNR	BD-Rate
0.79	**-14.99**



Subject quality comparison

- An artifact (propagation & blurring of chroma components) are visible especially when coded at low bit rates

- The artifact is removed by considering chroma components in the R-D calculation process at encoder side

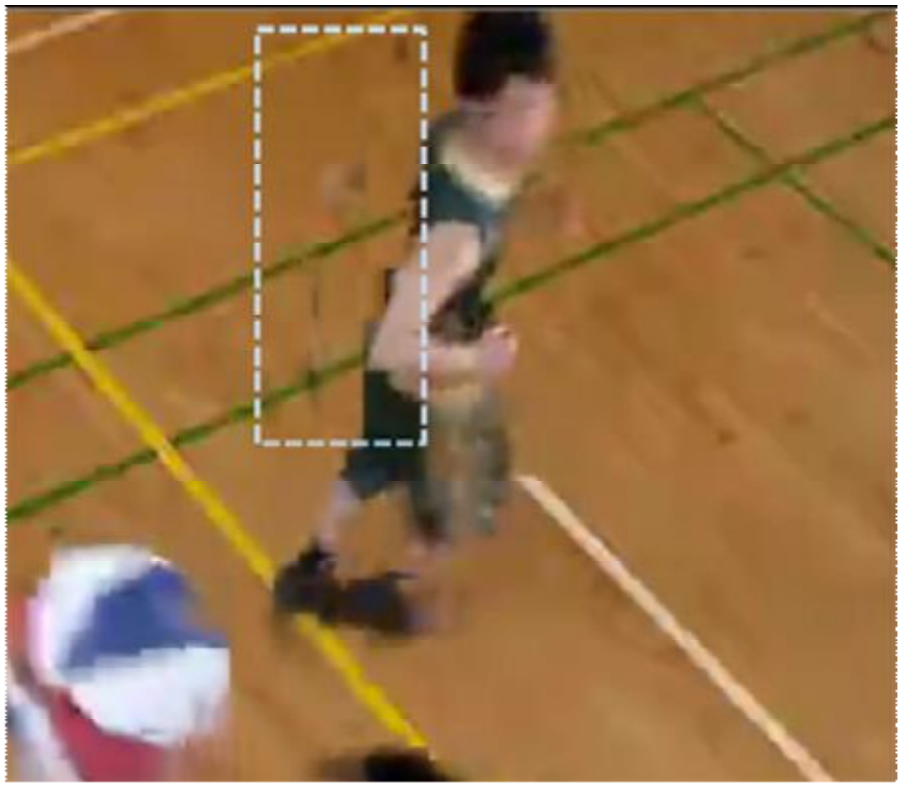


Alpha anchor

Proposed model

CS1 - Class C PartyScene @R2 (512kbps)

Subject quality comparison (Cont’d)



Alpha anchor

Proposed model

CS1 - Class C BasketballDrill @R2 (512kbps)

Subject quality comparison (Cont'd)

Alpha anchor

Proposed model

CS1 - Class C BasketballDrill @R2 (512kbps)

Software implementation

- On top of JM 11.0, new tools are implemented
- C programming language
- Complied using Microsoft Visual Studio 2008 under Microsoft Widows XP 32bit edition platform
- No code-level optimization for complexity

Complexity

- **Time Measurement**
  - Consumed time for encoding & decoding
  - Platform
    - 32bit executables are used in encoder and decoder.
    - CPU : Intel Core i7 950 @ 3.07Ghz (quad core)
    - Memory : 12 GB
  - Two target rates are tested due to time limitation
    - Rate2 (middle rate) / Rate5 (highest rate)

- **Encoder / Decoder complexity**
  - Encoder
    - CS1 : 9.27 times more than JM16.1 (4.64 times if implemented on the latest JM)
    - CS2 : 18.73 times more than JM16.1 (9.37 times if implemented on the latest JM)
  - Decoder
    - CS1 : 7.69 times more than JM16.1 (2.20 times if implemented on the latest JM)
    - CS2 : 7.11 times more than JM16.1 (2.03 times if implemented on the latest JM)
  - Comparison between JM11.0 and the latest JM
    - JM11.0 encoder is about 2 times slower than the latest JM
    - JM11.0 decoder is about 3.5 times slower than the latest JM

Conclusion

- LG's proposed model is based on AVC/H.264 JM 11.0
- Outperforms alpha anchor and gamma anchor, achieving 25.83% and 37.01% bit rate reduction under CS1 and CS2 respectively
- Complexity reduction was not considered in the submitted model.
- Coding performance is expected to be further improved if the additional tools are implemented (AWR, PAIF, MVC with B skip/direct, Chroma estimation mode with phase shift)