# Samsung and Vidyo response to CfP on Scalable Coding Extensions of HEVC

**Jill Boyce (Vidyo)**
**Ken McCann (Samsung)**

# Overview

Inspire the World, Create the Future


1 Introduction

2 Algorithm Description

3 Compression Performance

4 Complexity Analysis

5 Conclusions

# Introduction : Samsung/Vidyo Coding Framework

- This presentation covers JCTVC-K0044 and JCT-VC-K0045
- The Samsung/Vidyo coding framework provides the ability to trade off complexity and compression efficiency
  - Inter-layer and single-layer tools

**K0045** : low-complexity operating point focused on inter-layer tools, with good compression efficiency

- Average luma BD-rate gains of 30.3% / 47.1% for RA 2x / 1.5x
- Decoding time about ×1.8 relative to anchor

**K0044** : higher efficiency operating point, with a moderate increase in complexity

- Average luma BD-rate gains of 34.6% / 50.6% for RA 2x / 1.5x
- Decoding time about ×2.9 relative to anchor

# Introduction : Key Features

| Coding Tools | K0045 | K0044 | Notes |
|---|---|---|---|
| Intra_BL mode | √ | √ | Inter-layer Coding tool |
| Inter_BL mode | √ | √ | |
| Difference mode | √ | √ | |
| Inter-layer spatial MPM prediction | √ | √ | |
| Inter-layer motion prediction | √ | √ | |
| Inter-layer SAO | √ | √ | |
| Asymmetric motion partitions in enhancement layer | | √ | Main Profile Tool (HM8.0) |
| Bi-directional optical flow | | √ | Single-layer coding tool |
| Adaptive loop filter in enhancement layer | | √ | |
| TU-based implicit dQP refinement | | √ | |
| Multi-parameter probability estimation in enhancement layer CABAC | | √ | |

# Description: Units and Up-sampling

K0044 K0045

- **CTU/TU/PU**
  - Similar structure to non-scalable HEVC
  - PU is inferred to be 2Nx2N for some CU modes
    - Intra_BL
    - Inter_BL
- **Up-sampling filter**
  - Separable 2D FIR filter: 8 taps for Luma, 4 taps for Chroma
  - 1/4 phase shift :
    - the same DCT-IF as for HEVC motion compensation interpolation
  - 1/3 phase shift:
    - { 1, 4, -11, 52, 26, -8, 3, -1}/64
    - { -5, 50, 22,-3 }/64

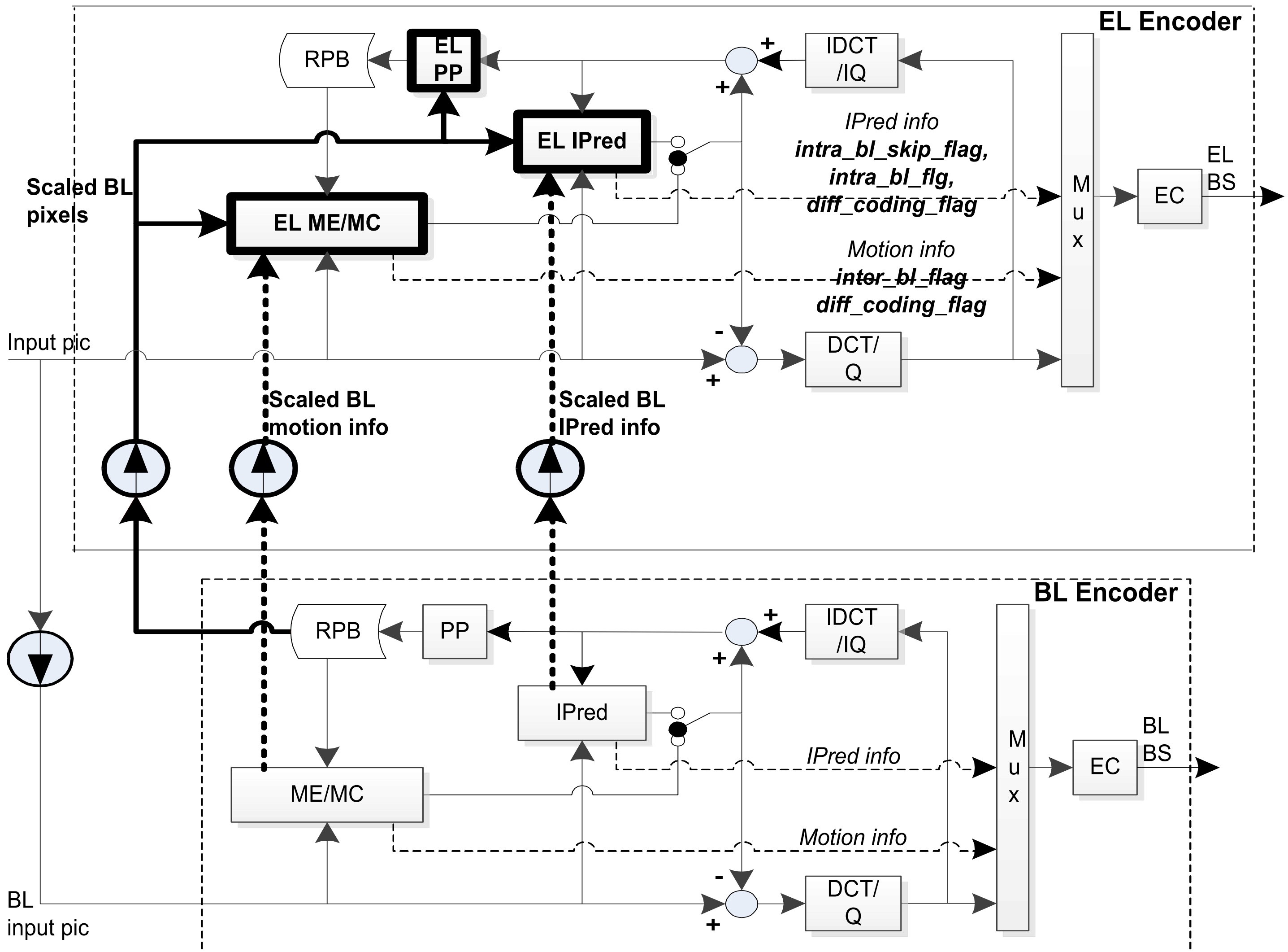
# Inter-Layer Prediction

K0044 K0045

- Prediction of Enhancement Layer (EL) from Base Layer (BL)
  - Double loop decoding
  - RD-optimized decision in encoder between various possible CU modes for EL
    - Intra_BL: BL pixels are scaled to EL CU resolution
    - Inter_BL: where BL motion vector is applied to EL
    - Pixel Coding: non-scalable coding of EL
    - Difference Coding: intra and inter prediction of EL incorporate the coded pixel values of the BL

# Proposed Scalable Encoder

K0044 K0045



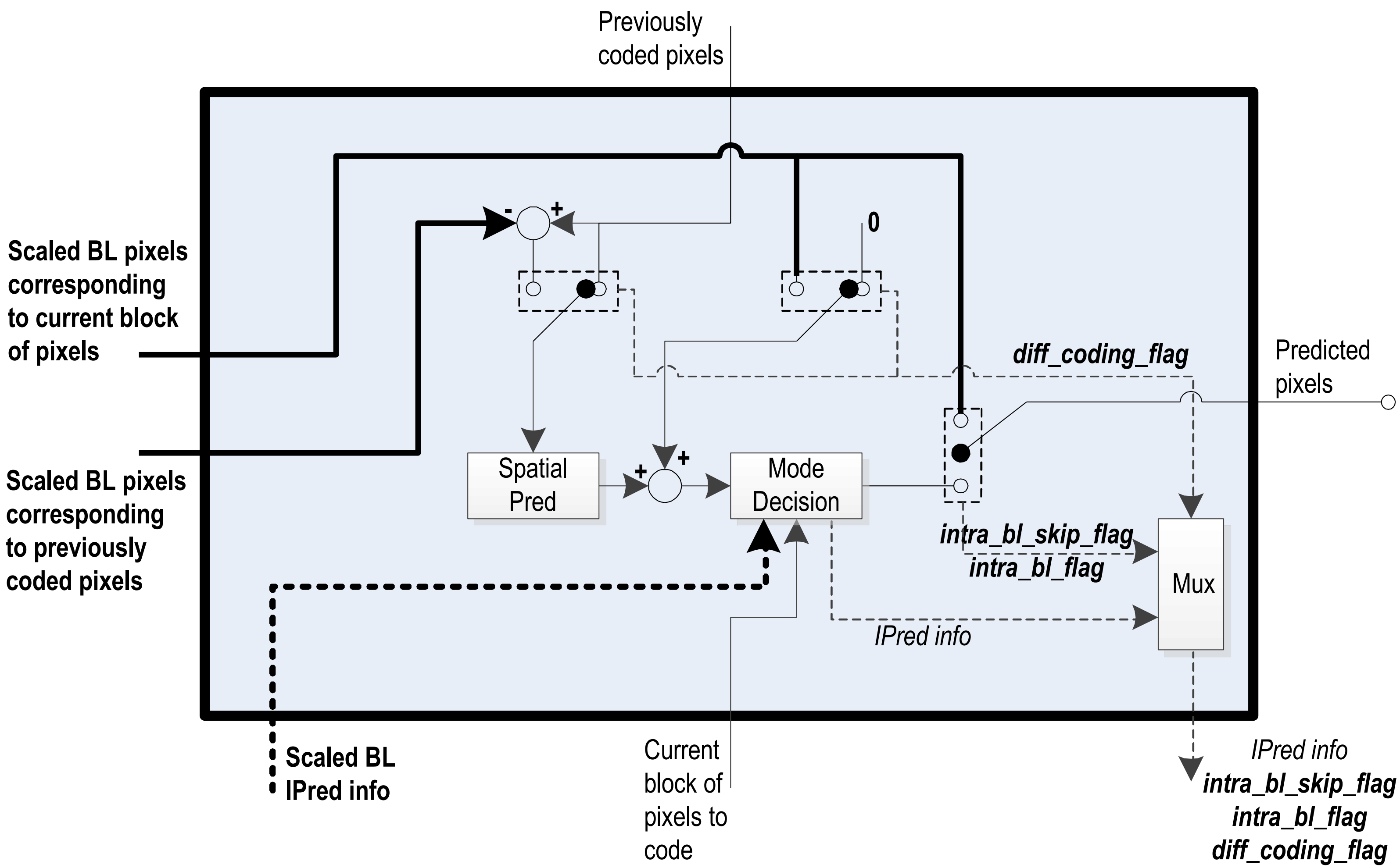
# EL Intra CU Coding

K0044 K0045

- **Intra_BL mode**
  - BL pixel block scaled to EL resolution and used as predictor for EL CU
  - Residual  coded by HEVC inter residual coding with prediction set to intra DC mode
  - Intra_BL Skip indicates that residual coefficients are zero
- **Difference Coding mode**
  - Combines intra prediction and inter-layer prediction
  - Spatial intra directional prediction performed using previously coded difference values
  - Add current BL pixel block scaled to EL resolution

# EL Intra Prediction Module

K0044 K0045



# EL Inter CU Coding

K0044 K0045

- **Inter_BL mode**
  - motion information for each 8x8 block of the CU is derived from corresponding BL block

- **Difference Coding mode**
  - Combines inter prediction and inter-layer prediction
  - Motion compensated blocks consist of difference values
  - Use bi-linear interpolation for sub-pel motion compensation
  - Add current BL pixel block scaled to EL resolution

# EL Coding Tools

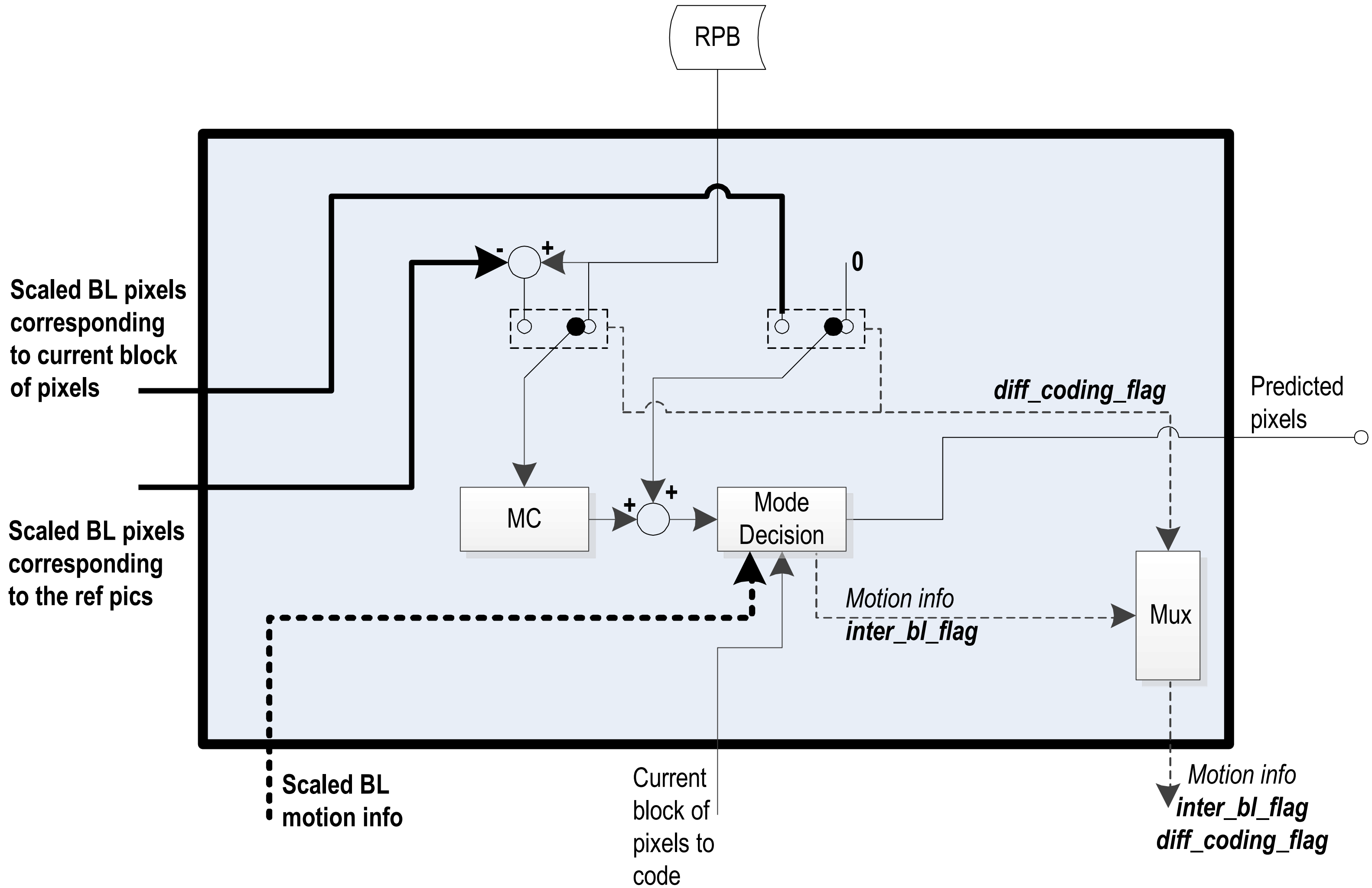
K0044 K0045

- **Motion vector prediction when Inter_BL mode not used**
  - For each PU, motion info from corresponding BL block is added as a candidate MV predictor, to the last position on the list
  - BL MV's stored for 4x4 blocks, EL selects position which corresponds to center of EL PU

- **Deblocking filter modification**
  - Additional condition to check for boundary between difference mode and non-difference mode blocks, set to Bs=1
  - Setting filter edges for Inter_BL mode are based on 8x8 PU boundaries as well as transform unit boundaries

# EL Inter Prediction Module

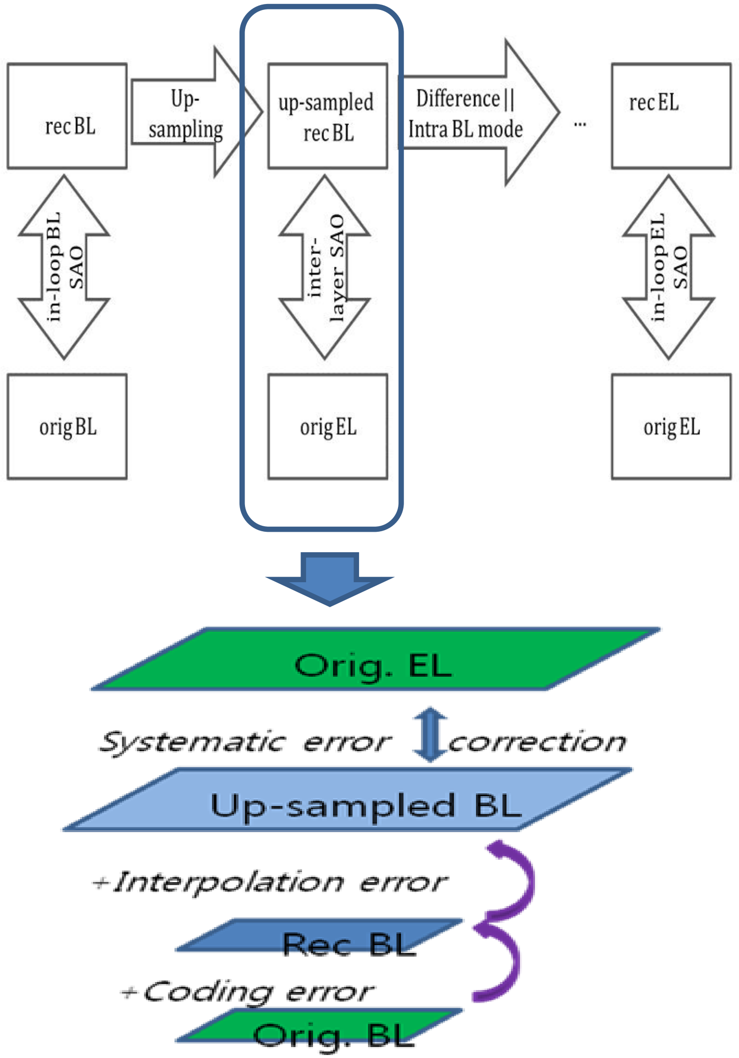
K0044 K0045



# Inter Layer SAO

K0044 K0045

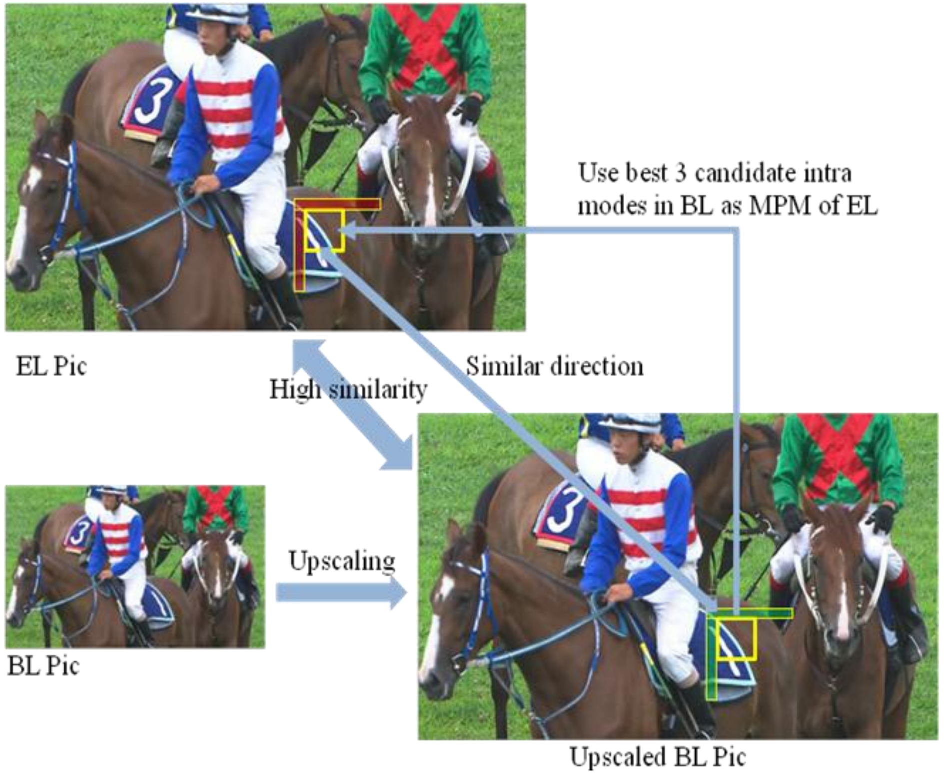
- **Re-use existing module:**
  - **Sample processing part is equal to in-loop SAO in non-scalable HEVC spec.**
- **No additional memory usage:**
  - **Re-use memory buffer reserved for in-loop SAO on EL**
- **Difference:**
  - **EO and BO are both applied**



# Inter Layer MPM Prediction (HE)



- Using high texture similarity between different resolution images.
- 3-MPMs of EL are set from up-scaled BL picture.



# Inter Layer MPM Prediction (LC)

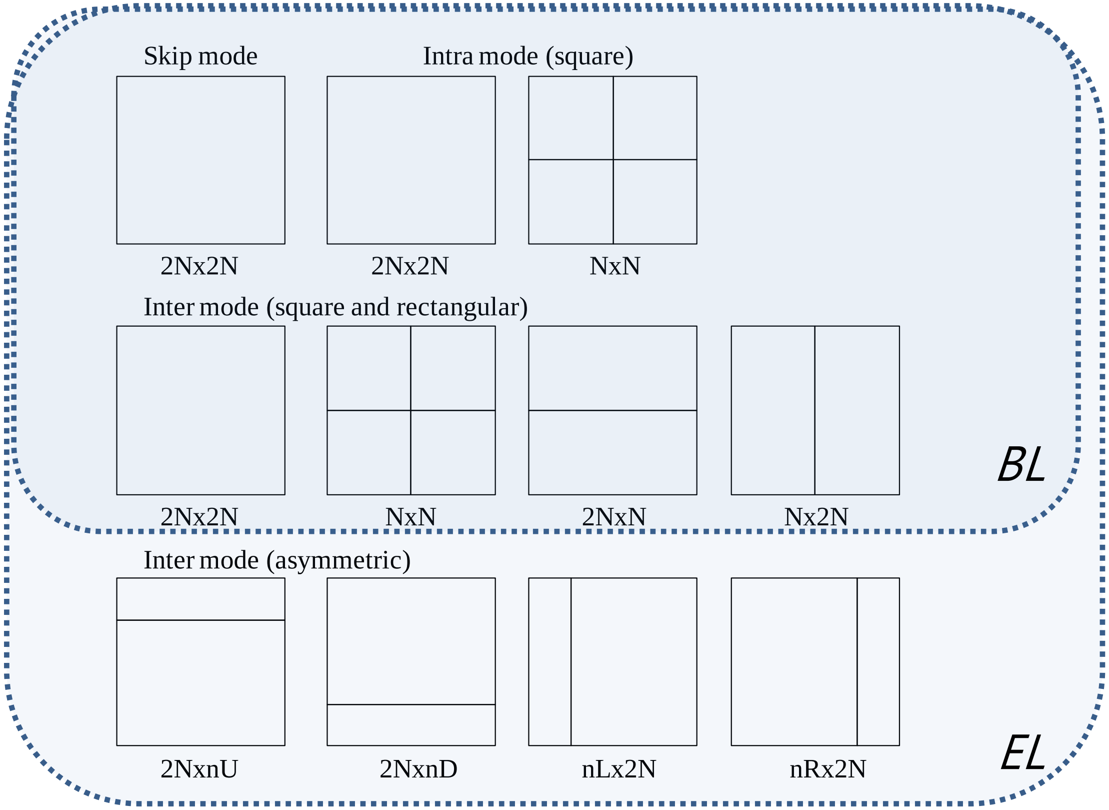


- The current non-scalable HEVC design's derivation process for MPMs yields three distinct MPMs
- For the EL PU, if the corresponding BL PU is coded in intra mode,
  - It is used as the very first MPM
  - If the base layer is the same with one of the two target layer MPMs
    - the base layer MPM is used as the first
    - the different target layer MPM is used as the second
    - the last third MPM derivation follows the process which is defined in the current HEVC for the case of two spatial intra modes having different intra prediction modes
  - Otherwise if the first two target layer MPMs that is different from the base layer MPM is used as the second and third MPMs, respectively

# AMP in EL

K0044

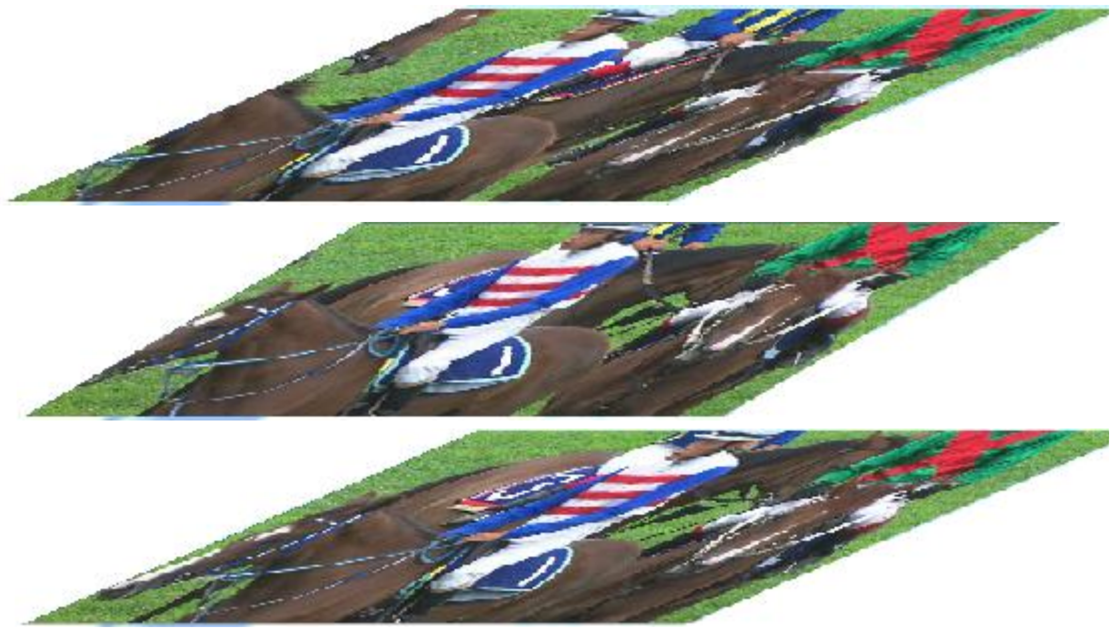
- **Only enable it on EL**
  - AMP is HEVC main profile tool

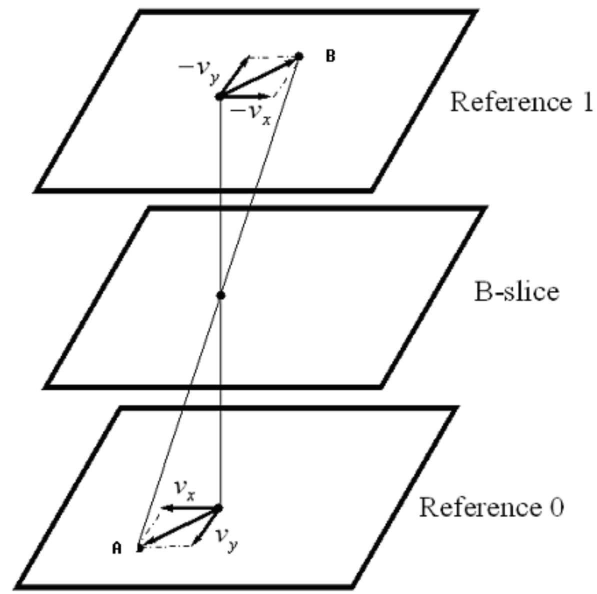


# Bi-Directional Optical flow

K0044

- **Pixel-wise compensation without additional syntax considerably increases coding efficiency for B-slices.**





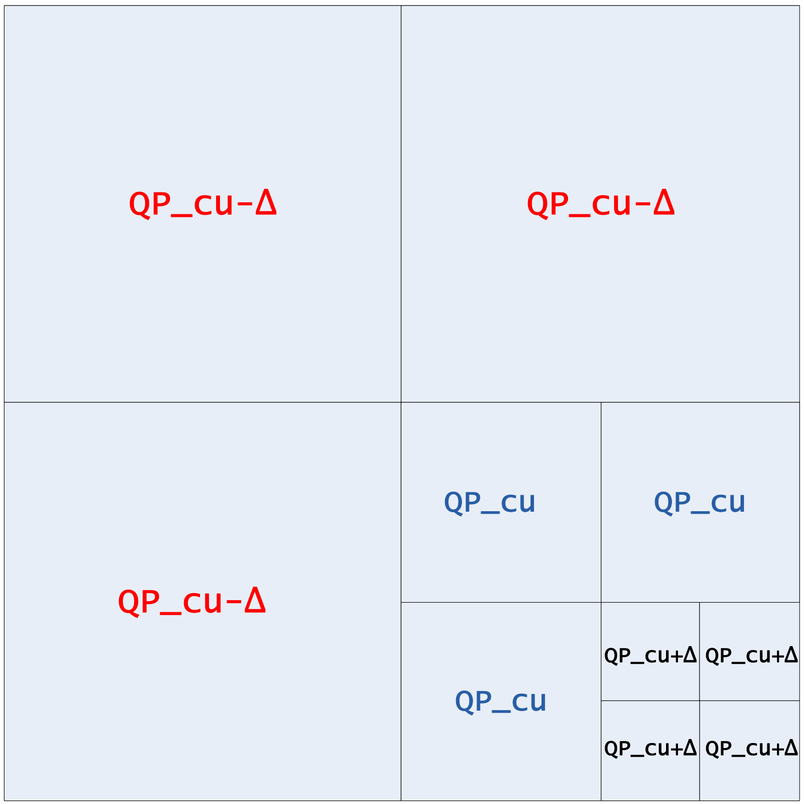

$$P_{OpticalFlow}[i,j] = \left(P^{(0)}[i,j] + P^{(1)}[i,j] + v_x[i,j]\left(I_x^{(0)}[i,j] - I_x^{(1)}[i,j]\right) + \right.$$
$$\left. + v_y[i,j]\left(I_y^{(0)}[i,j] - I_y^{(1)}[i,j]\right) + 1\right) >> 1$$

# TU size based implicit dQP refinement

K0044

CU = 64

QP_cu = QP +dQP




- **Implicit signaling of dQP based on transform size (TU based dQP)**
- **The static areas have more possibility to be reference rather than moving region.**
- **Relative higher QP is applied for small transform block size.**

| | 4x4 | 8x8 | 16x16 | 32x32 |
|---|---|---|---|---|
| Implicit dQP | 2xΔ | Δ | 0 | -Δ |

## Enable ALF on EL

- ALF design follows HM 5.0 with minor modification:
- Block based adaptation with directional features classification
  - *Pixel directional features:*
    - $H(i,j) = | X(i,j)<<1 - X(i,j-1) - X(i,j+1) |$
    - $V(i,j) = | X(i,j)<<1 - X(i-1,j) - X(i+1,j) |$
  - block-wise features :
    - $H_B = \sum_{i=1,2} \sum_{j=1,2} H(i,j)$
    - $V_B = \sum_{i=1,2} \sum_{j=1,2} V(i,j)$
    - $L_B = (H_B + V_B) >> 2$

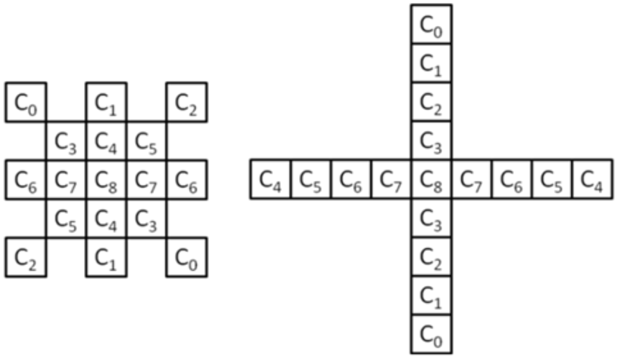
$$dir = \begin{cases} 1, \text{ if } V_B > 2 \times H_B \\ 2, \text{ if } H_B > 2 \times V_B \\ 0, \text{ otherwise} \end{cases}$$

$$norMag = \max\{ 7, L_B >> 1 \}$$

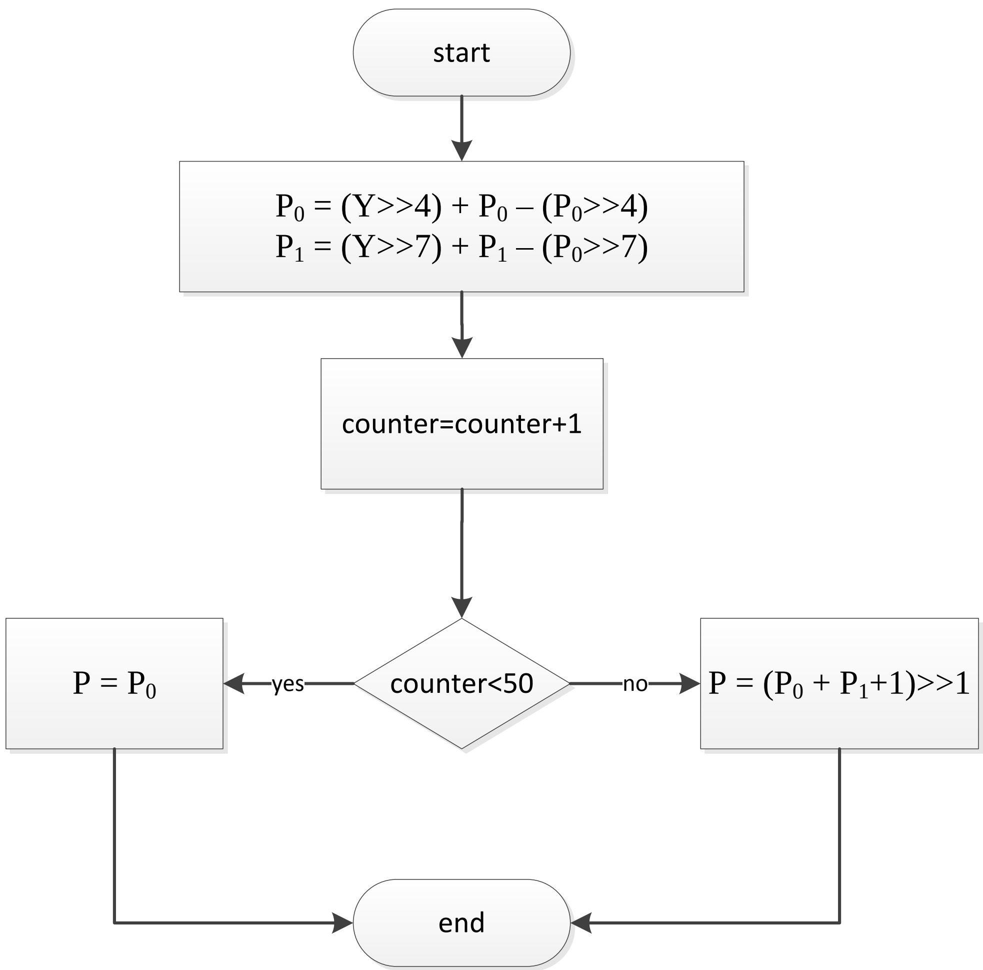
*Class label* $C_B$ = **classTab** [*dir*][*norMag*]

| | norMag | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | ***0*** | ***1*** | ***2*** | ***3*** | ***4*** | ***5*** | ***6*** | ***7*** |
| ***dir=0*** | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| ***dir=1*** | 0 | 3 | 3 | 3 | 3 | 3 | 3 | 4 |
| ***dir=2*** | 0 | 5 | 5 | 5 | 5 | 5 | 5 | 6 |

- Filter shapes:
  - snowflake-5x5
  - cross-9x9



# Multi-parameter probability estimation in EL K0044




- **Multi-parameter probability estimation for CABAC in enhancement layer used the technique proposed in JCTVC-F254**

# Compression Performance (1) : JCTVC-K0045

## 1. Spatial scalability

| Random Access HEVC 2x | | | Random Access HEVC 1.5x | | |
|---|---|---|---|---|---|
| Y | U | V | Y | U | V |
| -32.7% | -24.3% | -25.3% | | | |
| -27.1% | -21.0% | -19.4% | -46.9% | -41.4% | -40.3% |
| **-29.9%** | **-22.7%** | **-22.4%** | **-46.9%** | **-41.4%** | **-40.3%** |

| All Intra HEVC 2x | | | All Intra HEVC 1.5x | | |
|---|---|---|---|---|---|
| Y | U | V | Y | U | V |
| -39.1% | -37.1% | -37.9% | | | |
| -31.6% | -29.3% | -29.2% | -54.1% | -52.2% | -52.4% |
| **-35.3%** | **-33.2%** | **-33.5%** | **-54.1%** | **-52.2%** | **-52.4%** |

<Target Rate>

| | Random Access HEVC 2x | | | Random Access HEVC 1.5x | | |
|---|---|---|---|---|---|---|
| | Y | U | V | Y | U | V |
| Class A+ | -33.3% | -25.0% | -26.0% | | | |
| Class B | -27.3% | -21.2% | -19.7% | -47.1% | -47.1% | -40.6% |
| Overall | **-30.3%** | **-23.1%** | **-22.8%** | **-47.1%** | **-47.1%** | **-40.6%** |

| | All Intra HEVC 2x | | | All Intra HEVC 1.5x | | |
|---|---|---|---|---|---|---|
| | Y | U | V | Y | U | V |
| Class A+ | -39.1% | -37.2% | -38.0% | | | |
| Class B | -31.6% | -29.4% | -29.2% | -54.1% | -52.3% | -52.4% |
| Overall | **-35.4%** | **-33.3%** | **-33.6%** | **-54.1%** | **-52.3%** | **-52.4%** |

<Actual Rate>

## 2. SNR scalability

| | Actual rate | | | Target rate | | |
|---|---|---|---|---|---|---|
| | Y | U | V | Y | U | V |
| Class A+ | -38.7% | -32.5% | -32.5% | -38.1% | -31.8% | -21.9% |
| Class B | -35.0% | -27.6% | -24.5% | -34.8% | -27.3% | -24.3% |
| Overall | **-36.8%** | **-30.0%** | **-28.5%** | **-36.4%** | **-29.6%** | **-28.1%** |

## 3. AVC scalability

| Hybrid vs HEVC 2x | | | Hybrid vs HEVC 1.5x | | |
|---|---|---|---|---|---|
| Y | U | V | Y | U | V |
| -31.9% | -23.1% | -24.1% | | | |
| -25.3% | -19.6% | -17.6% | -43.0% | -37.8% | -35.9% |
| **-28.6%** | **-21.4%** | **-20.8%** | **-43.0%** | **-37.8%** | **-35.9%** |

| | Hybrid vs HEVC 2x | | | Hybrid vs HEVC 1.5x | | |
|---|---|---|---|---|---|---|
| | Y | U | V | Y | U | V |
| Class A+ | -32.3% | -23.6% | -24.5% | | | |
| Class B | -25.5% | -19.9% | -17.9% | -43.2% | -38.0% | -36.2% |
| Overall | **-28.9%** | **-21.7%** | **-21.2%** | **-43.2%** | **-38.0%** | **-36.2%** |

# Compression Performance (2) : JCTVC-K0044

## 1. Spatial scalability

| Random Access HEVC 2x | | | Random Access HEVC 1.5x | | |
|---|---|---|---|---|---|
| Y | U | V | Y | U | V |
| -37.0% | -25.6% | -26.3% | | | |
| -31.5% | -22.0% | -19.6% | -50.3% | -42.4% | -40.3% |
| **-34.2%** | **-23.8%** | **-22.9%** | **-50.3%** | **-42.4%** | **-40.3%** |

| All Intra HEVC 2x | | | All Intra HEVC 1.5x | | |
|---|---|---|---|---|---|
| Y | U | V | Y | U | V |
| -40.6% | -38.9% | -39.7% | | | |
| -33.1% | -30.4% | -29.9% | -55.2% | -52.9% | -52.9% |
| **-36.9%** | **-34.6%** | **-34.8%** | **-55.2%** | **-52.9%** | **-52.9%** |

<Target Rate>

| | Random Access HEVC 2x | | | Random Access HEVC 1.5x | | |
|---|---|---|---|---|---|---|
| | Y | U | V | Y | U | V |
| Class A+ | -37.4% | -26.1% | -26.8% | | | |
| Class B | -31.7% | -22.3% | -19.8% | -50.5% | -42.7% | -40.6% |
| **Overall** | **-34.5%** | **-24.2%** | **-23.3%** | **-50.5%** | **-42.7%** | **-40.6%** |

| | All Intra HEVC 2x | | | All Intra HEVC 1.5x | | |
|---|---|---|---|---|---|---|
| | Y | U | V | Y | U | V |
| Class A+ | -40.6% | -39.0% | -39.8% | | | |
| Class B | -33.2% | -30.4% | -29.9% | -55.2% | -53.0% | -52.9% |
| Overall | **-36.9%** | **-34.7%** | **-34.8%** | **-55.2%** | **-53.0%** | **-52.9%** |

<Actual Rate>

## 2. SNR scalability

| | Actual rate | | | Target rate | | |
|---|---|---|---|---|---|---|
| | Y | U | V | Y | U | V |
| Class A+ | -43.2% | -33.0% | -32.7% | -42.7% | -32.3% | -32.1% |
| Class B | -39.1% | -28.6% | -24.2% | -38.8% | -28.2% | -23.8% |
| Overall | **-41.2%** | **-30.8%** | **-28.5%** | **-40.7%** | **-30.3%** | **-28.0%** |

## 3. AVC scalability

| Hybrid vs HEVC 2x | | | Hybrid vs HEVC 1.5x | | |
|---|---|---|---|---|---|
| Y | U | V | Y | U | V |
| -36.0% | -24.2% | -24.8% | | | |
| -30.0% | -20.6% | -17.4% | -46.8% | -38.7% | -35.8% |
| **-33.0%** | **-22.4%** | **-21.1%** | **-46.8%** | **-38.7%** | **-35.8%** |

| | Hybrid vs HEVC 2x | | | Hybrid vs HEVC 1.5x | | |
|---|---|---|---|---|---|---|
| | Y | U | V | Y | U | V |
| Class A+ | -36.5% | -24.8% | -25.3% | | | |
| Class B | -30.2% | -20.9% | -17.7% | -47.1% | -39.0% | -36.1% |
| Overall | **-33.4%** | **-22.8%** | **-21.5%** | **-47.1%** | **-39.0%** | **-36.1%** |

# Complexity Analysis

- **Run time increment relative to HM6.1**

| test | K0044 Enc | K0044 Dec | K0045 Enc | K0045 Dec |
|---|---|---|---|---|
| RA HEVC spatial | ×4.1 | ×2.9 | ×2.8 | ×1.8 |
| AI HEVC | ×3.5 | ×2.4 | ×2.8 | ×1.6 |
| RA HEVC SNR | ×4.9 | ×3.3 | ×3.4 | ×2.4 |
| Avg. | **×4.2** | **×2.9** | **×3.0** | **×1.9** |

# Conclusions

- The Samsung/Vidyo coding framework has been described in some detail
- CfP responses submitted in all test categories
  - including BL encoded with AVC
- Our responses to the CfP demonstrated two key operating points
  - K0045: low-complexity operating point, with good compression efficiency
  - K0044: higher efficiency operating point, with a moderate increase in complexity
- We propose this coding framework as a candidate for the Test Model to facilitate further work

Inspire the World, Create the Future

# THANK YOU