Master: INFORMATIQUE
Parcours: VICO Visual Computing

UE: Multimedia Communication

video coding

vincent.ricordel@univ-nantes.fr

References and Text books

- Text book: M. Wien: *High Efficiency Video Coding – Coding Tools and Specification*. Springer 2014 -> not available for free!

- Reference Paper: Sullivan, Gary J., et al. "*Overview of the high efficiency video coding (HEVC) standard*." Circuits and Systems for Video Technology, IEEE Transactions on22.12 (2012): 1649-1668 -> complete and detailed

- Need to read further? scholar.google.com is the best option

- **Course material will be available in *ExtraDoc***

Contents

- Introduction
- Basic Concepts
- HEVC Overview
- Extension of HEVC
- HEVC Test Model (HM)

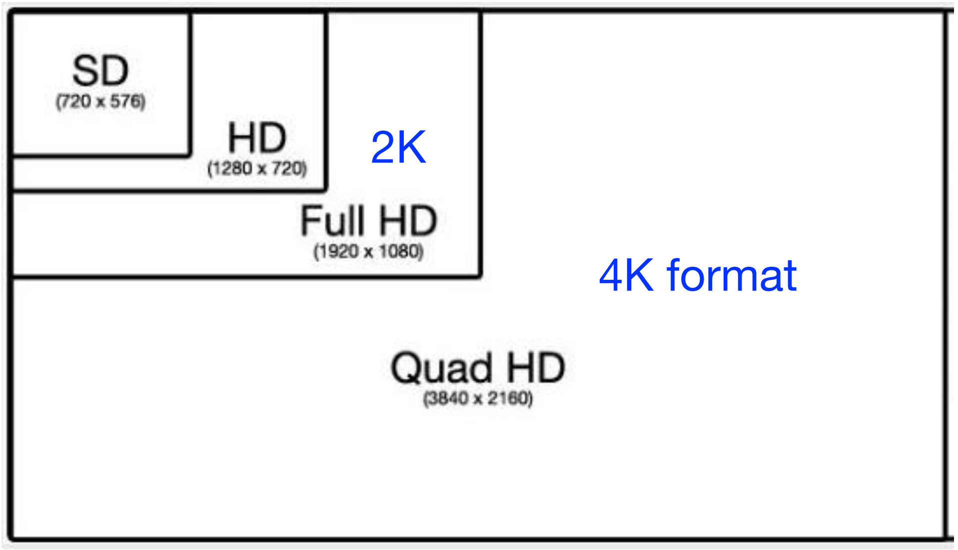
Introduction

Needs for *Efficient* Video Compression

- Limited Bandwidth
  Increasing demand on high quality video streaming (ex. Online movies, video conferencing, etc)

- Limited storage
  Need to store a huge amount of data on a limited memory (Ex. smart phone)

Needs for HEVC



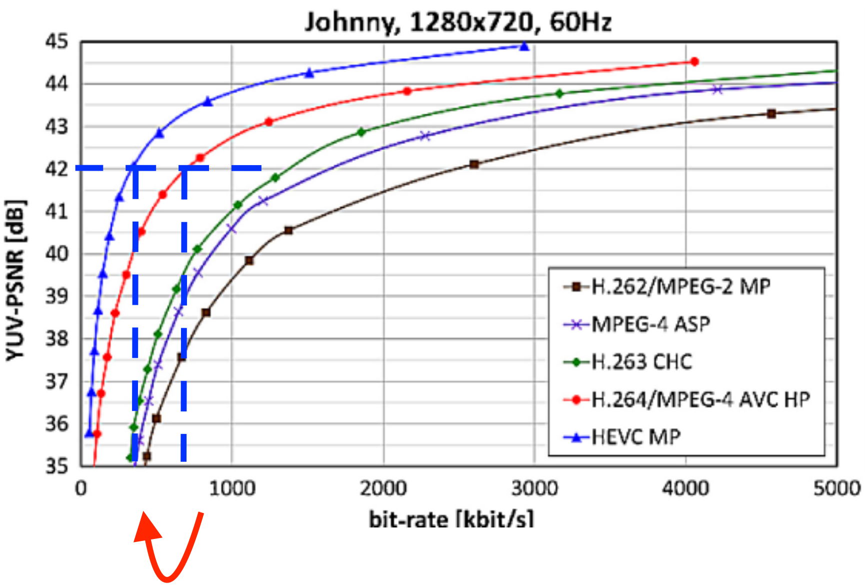
- Can videos be deployed with approximately the same bitrate when the size is doubled ? => HEVC instead of AVC

How to achieve 50% gain?

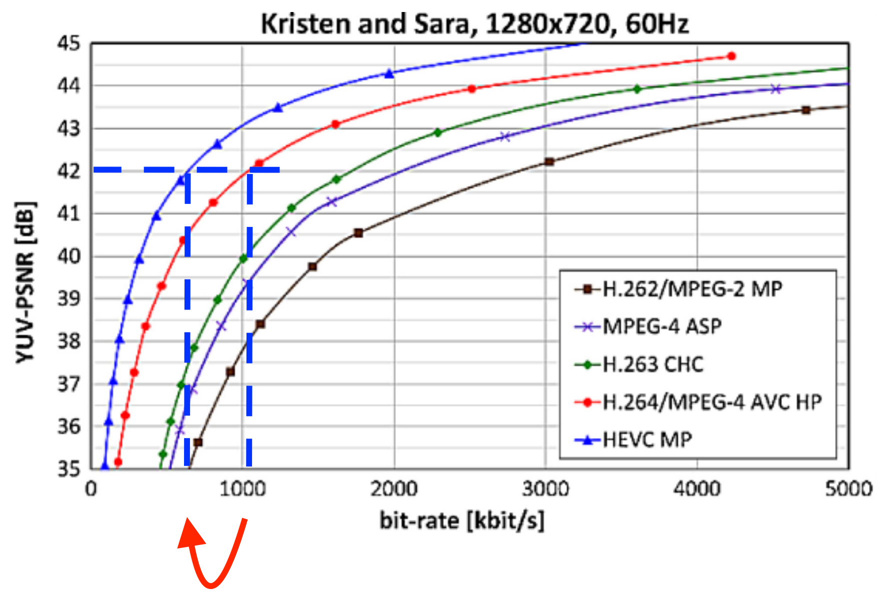
- Larger block sizes with flexible partitioning
- More intra prediction directions
- Better motion estimation (asymmetric partitioning + sub sample accuracy)

- Better reconstruction filters (deblocking + in loop filtering)

HEVC vs Other MPEG Standards



AVC -> HEVC



AVC -> HEVC

Ohm, J., et al. "Comparison of the coding efficiency of video coding standards—including high efficiency video coding (HEVC). " *Circuits and Systems for Video Technology, IEEE Transactions on* 22.12 (2012): 1669-1684.

Basic Concepts

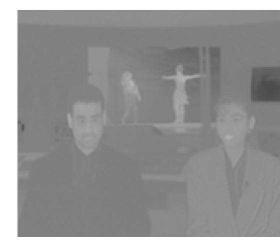
Must know!

- Image Representation (rgb, yuv)
  yuv420

u

Original

y



v

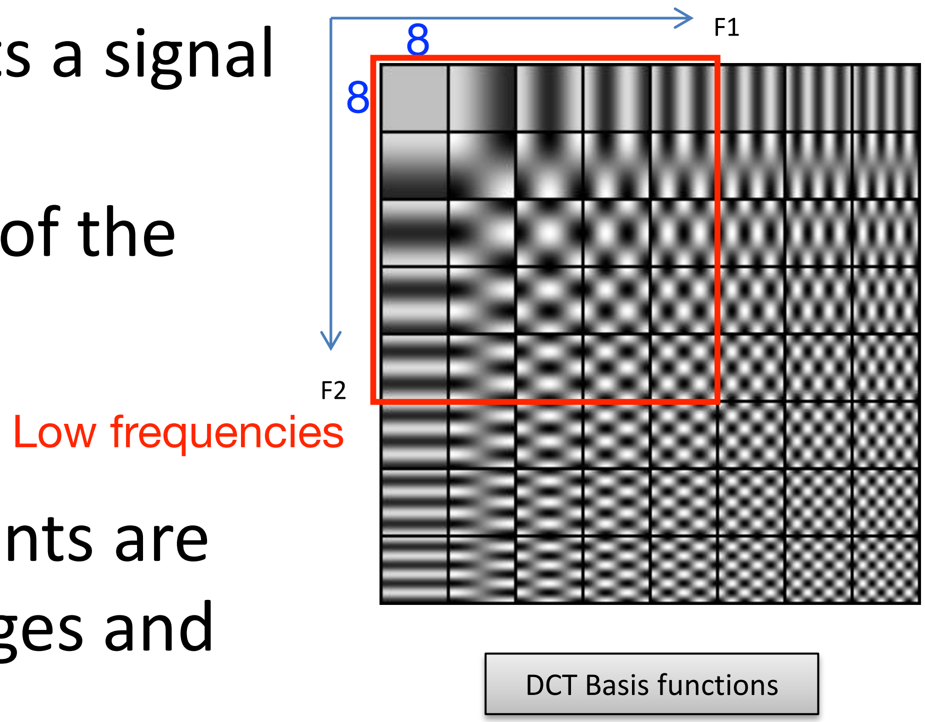
Must know

- Prediction (Intra, inter)
  Used to reduce redundancy in the spatial and temporal domain (resp.) => residual signal

- Image Transforms (DCT)
  further reduces the redundancy and arrange the coefficients for proper scanning (zigzag scanning)

Must know – 2D DCT

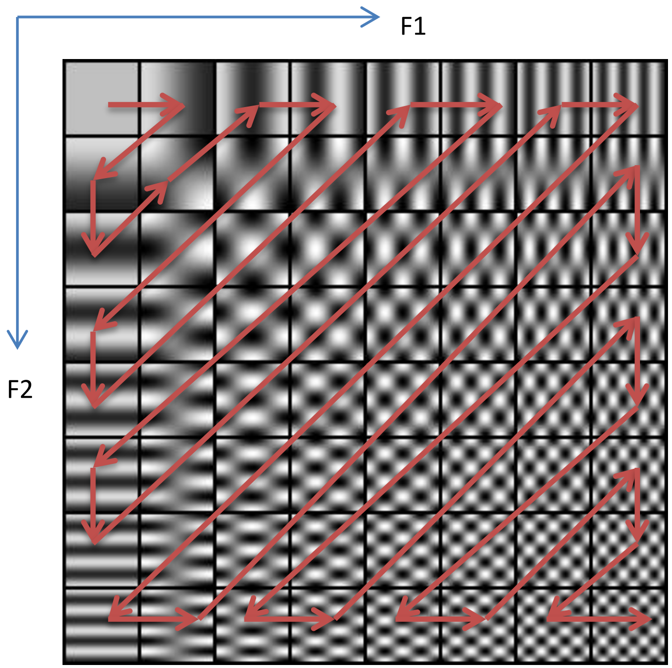
- DCT transform represents a signal with a set of coefficients representing the weight of the DCT basis functions
- Low frequency components are dominant in natural images and videos



Must know – Zigzag Scan

- Scanning the coefficients according to their importance:

  low frequencies (upper left part) are more important than high frequencies (lower right)



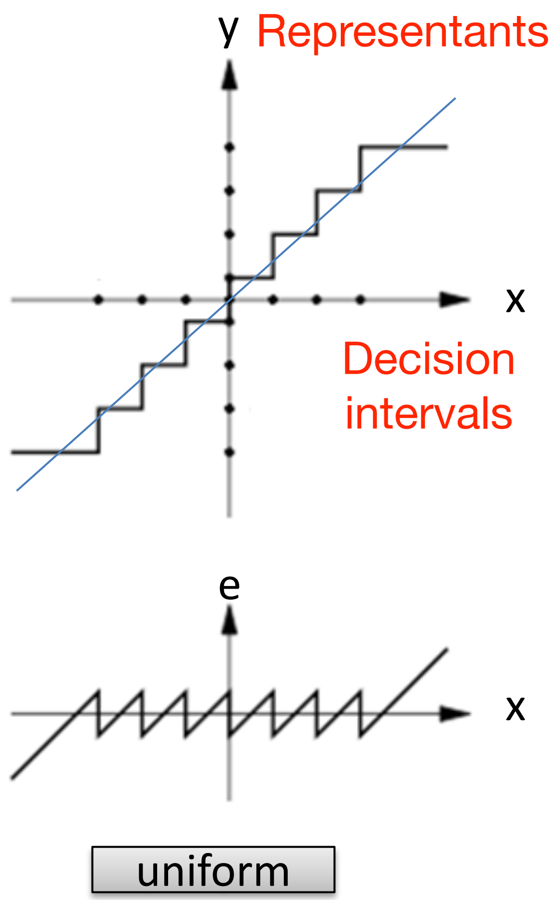
Must know – Quantization

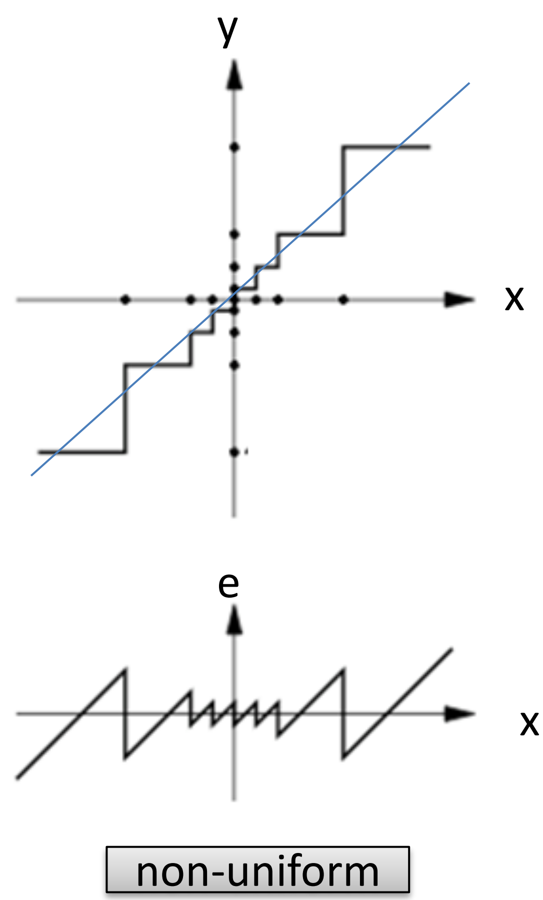
Input x is quantized to y

$y = Q(x)$

Quantization error

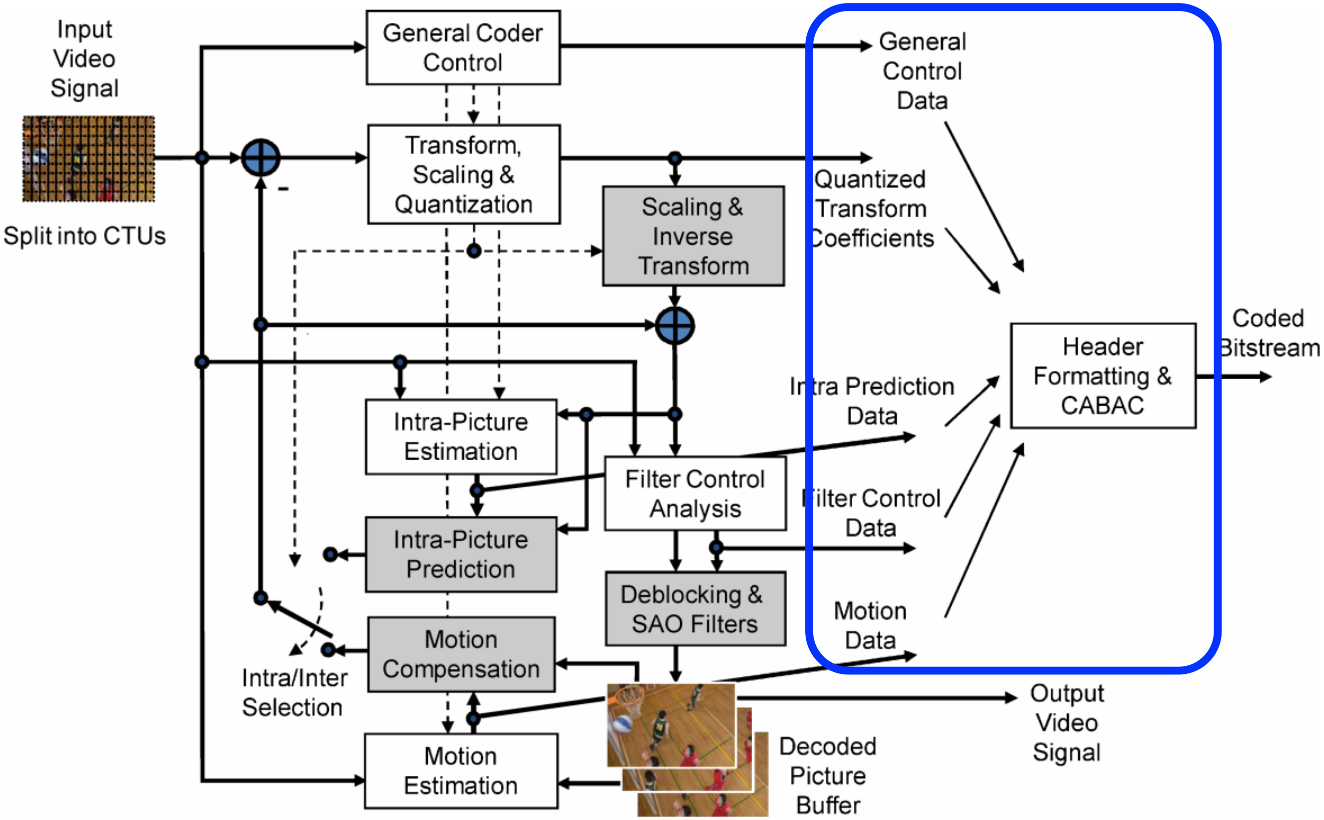
$e = y - x$





HEVC Overview

HEVC Reference Encoder



Sullivan, Gary J., et al. "Overview of the high efficiency video coding (HEVC) standard." *Circuits and Systems for Video Technology, IEEE Transactions on*22.12 (2012): 1649-1668.

Processing Units and Blocks

- Unit vs Block
  - Units contains data of the 3 components (y,u and v)
  - Blocks contains data of 1 specific components (y or u or v)

Texte

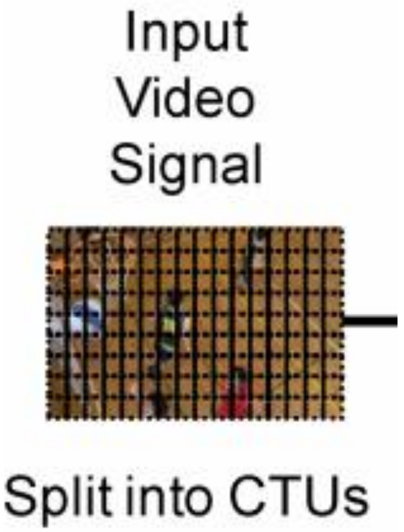
- Ex:
  - 16x16 luma block (y only)
  - 8x8 unit (8x8 luma + 2* 4x4 chroma blocks)

Assuming yuv420

Coding Tree Unit (CTU)

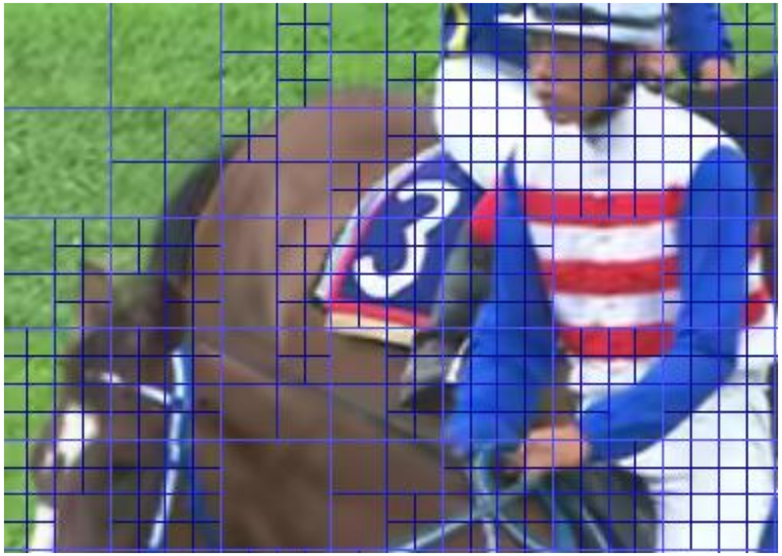
- Analogous to macro blocks in AVC
- Split the input videos into equal units
- Contains 3 Coding Tree Blocks (CTB)
  - LxL luma CTB
  - 2 (L/2)x(L/2) chroma CTB

  Assuming yuv420 format
- L can be 64, 32 and 16

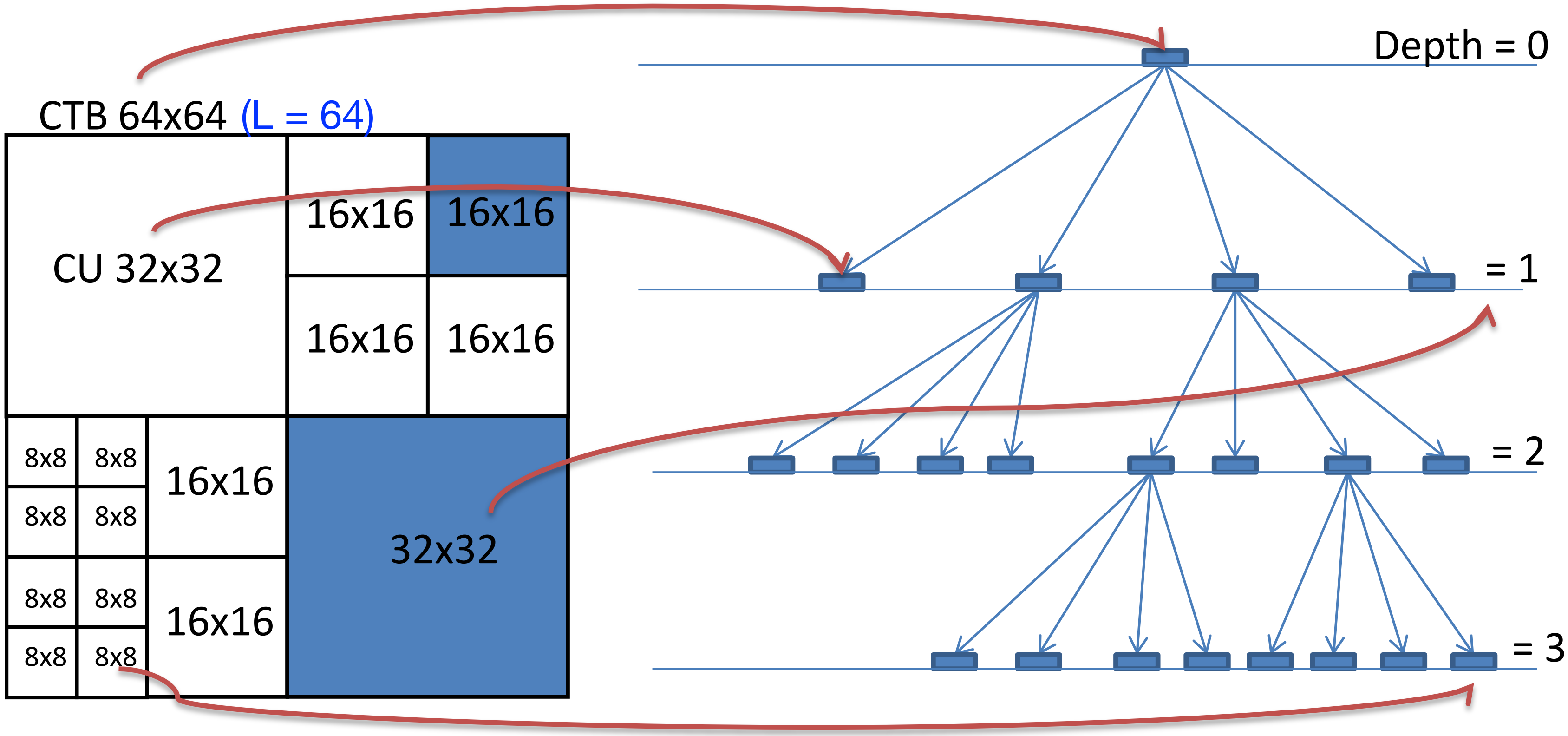


Coding Unit (CU)

- Each CTU starts with 1 CU
- CU can be split into 4 CUs in *Quad-Tree* manner
- CU contains Prediction Units (PUs) and Transform Units (TUs)

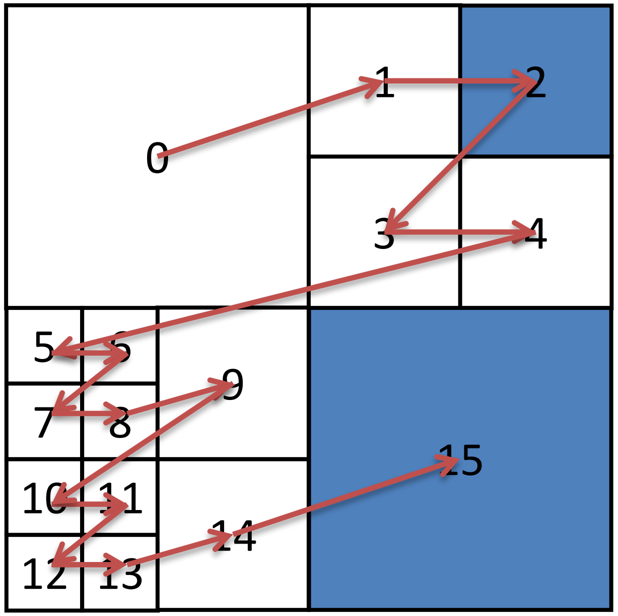


CU Splitting Depth



Maximum depth is 4

Z-Scan of CB



By following the leaves of the CTB

HEVC CTU quadtree partitioning

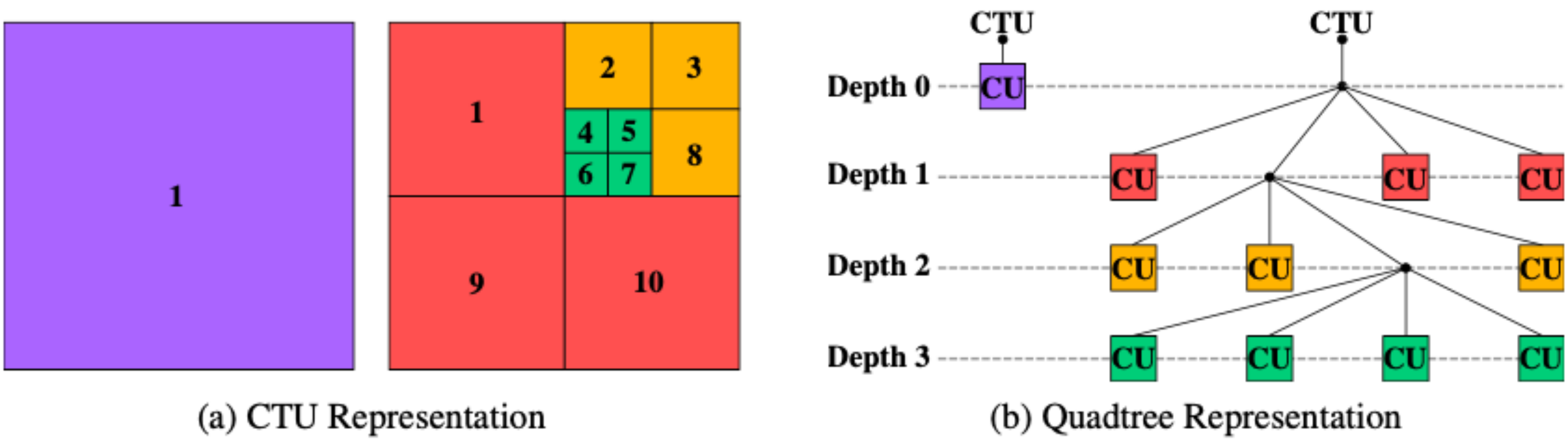


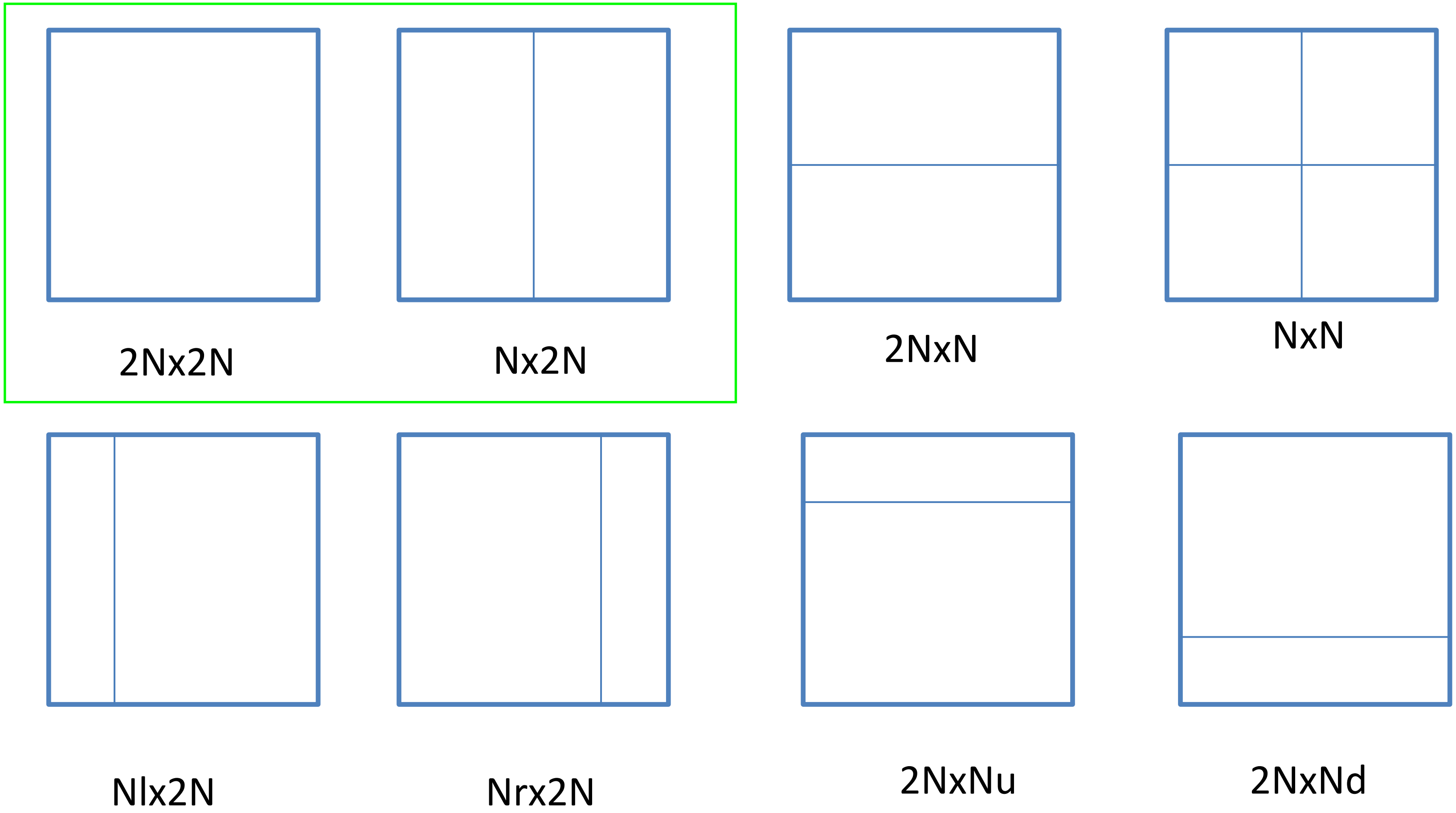
Fig. 2.8 Example of HEVC CTU quadtree partitioning.

Prediction Unit (PU)

- A basic unit in which the prediction is performed (intra / inter)

- Each CB can contain 1 or multiple PB according to the *mode splitting*

PU Mode Splitting



In intra prediction, only 2Nx2N and NxN is used!

HEVC PU prediction unit modes

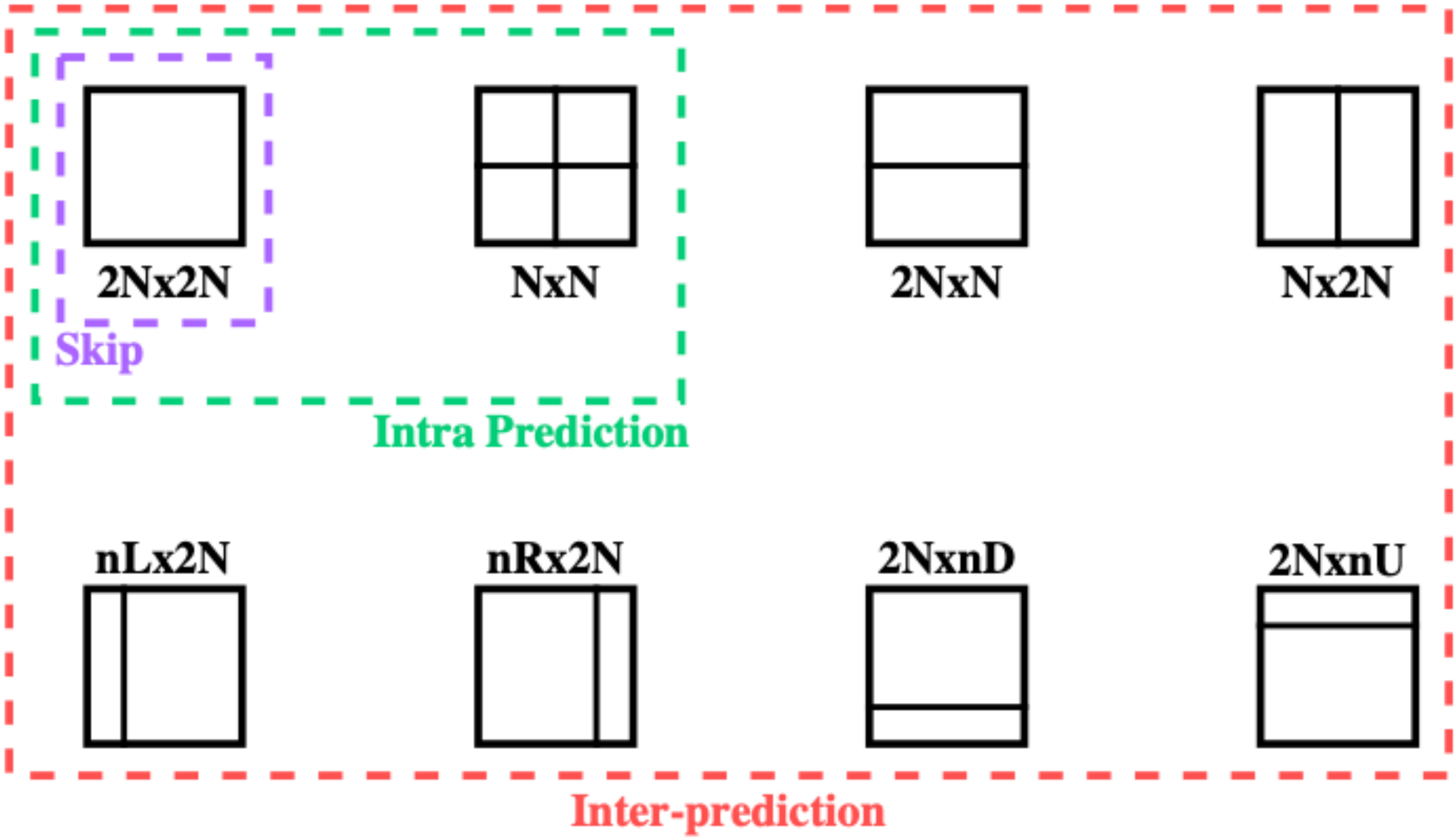


Fig. 2.9 HEVC Prediction Unit (PU) modes.

Why is it useful ?

Best prediction can be obtained if asymmetric partitioning is used!

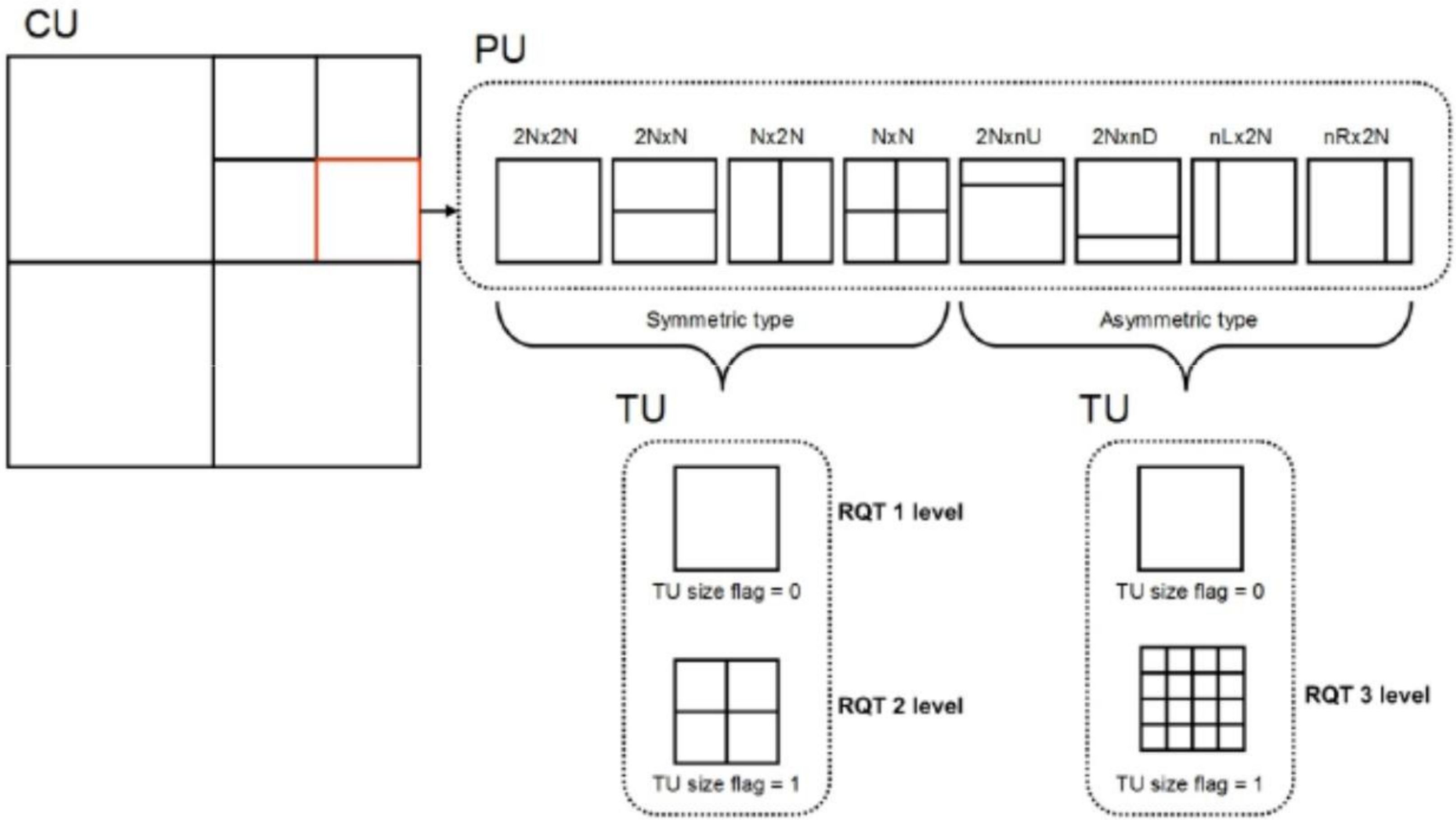
Transform Unit (TU)

- A basic unit where the transform coding to the residual signal is performed

- The residual signal can be also further split => residual quad tree RQT
- DCT is applied for each TB, DST is applied to TBs of intra-predicted luma blocks of size 4x4

DCT : Discrete Cosine Transform
DST : Discrete Sinus Transform

Putting everything Together

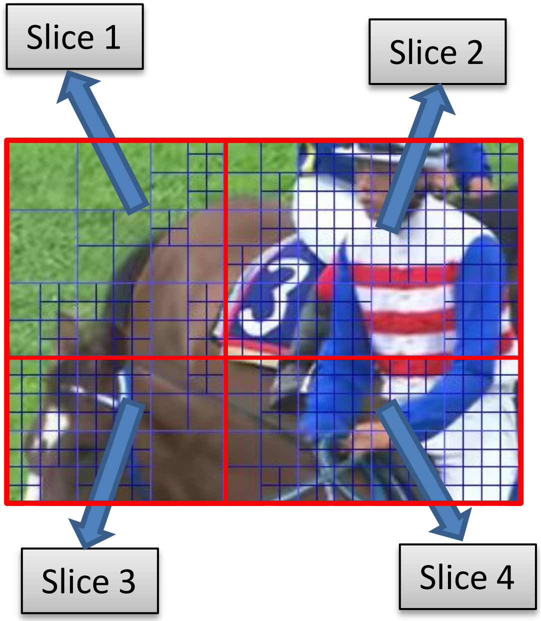


Slices

- Slices are sub-images that contains integer number of CTUs

- Independently decoded from the others in the same frame

- In HM, each slice is one frame!



HM: HEVC Test Model (the reference software for HEVC)

Slices

- Each slice is independently decodable
- Types of slices
  1. I Slice -> only intra prediction is allowed
  2. P Slice -> intra + inter (one reference) *
  3. B Slice -> intra + inter (two references) **

- In case of 1 slice / frame (ex. HM), we consider I-Frame, P-Frame and B-Frame

* P stands for predicted frame

** B stands for bi-predicted frame

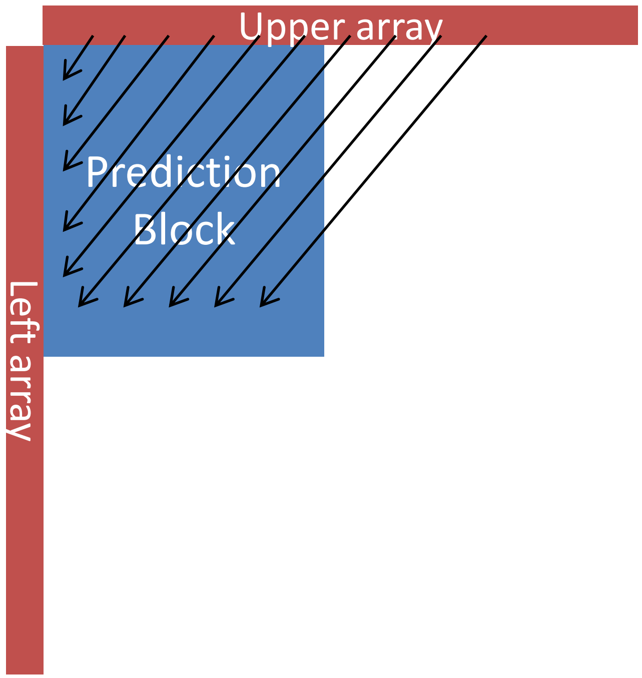
Prediction in HEVC

- Intra-Picture Prediction
  - Directional
  - DC
  - Planar

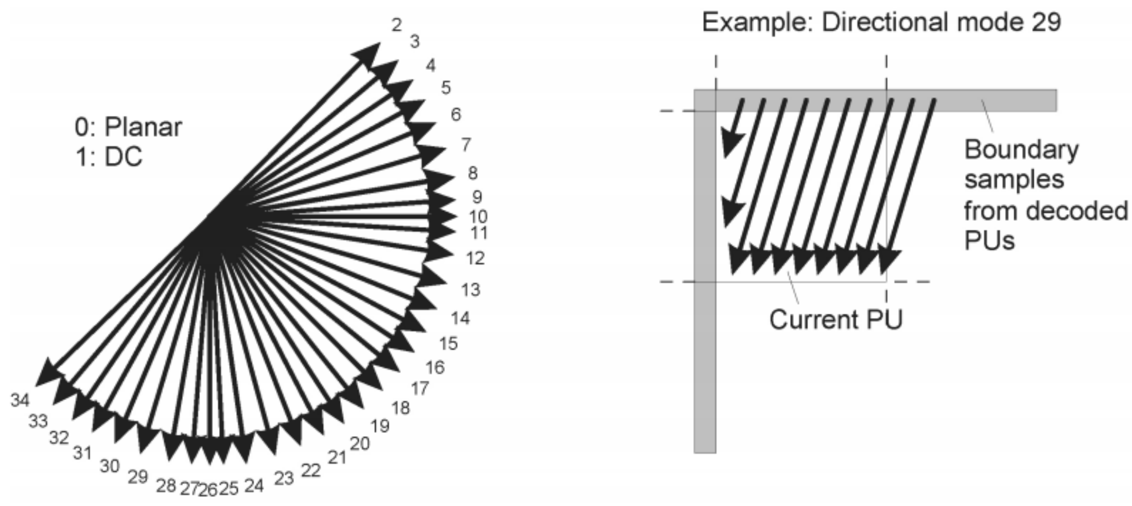
- Inter-Picture Prediction
  - Motion Compensation
  - Advanced Motion Vector Prediction

Intra-Picture Prediction

- HEVC uses upper and left elements to generate the prediction signal
- 33 Prediction direction + DC prediction + Planar prediction



Intra-Picture Prediction

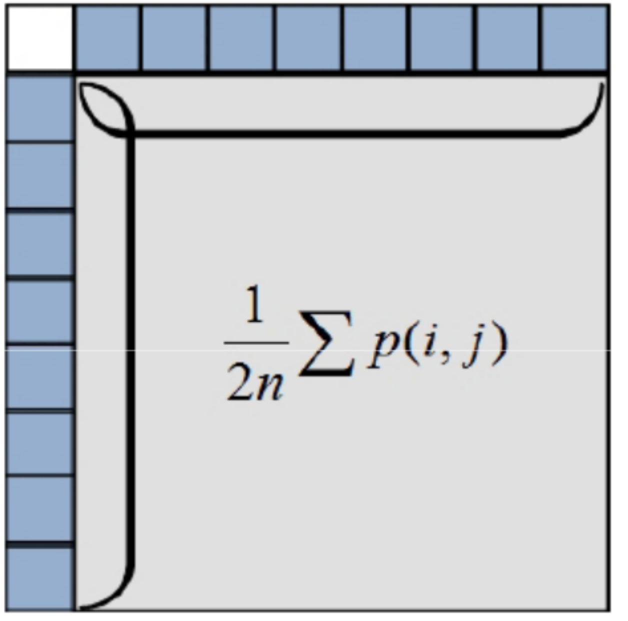


Sullivan, Gary J., et al. "Overview of the high efficiency video coding (HEVC) standard." *Circuits and Systems for Video Technology, IEEE Transactions on*22.12 (2012): 1649-1668.

DC and Planar Prediction

- DC prediction
  prediction signal equals to the average of all prediction elements

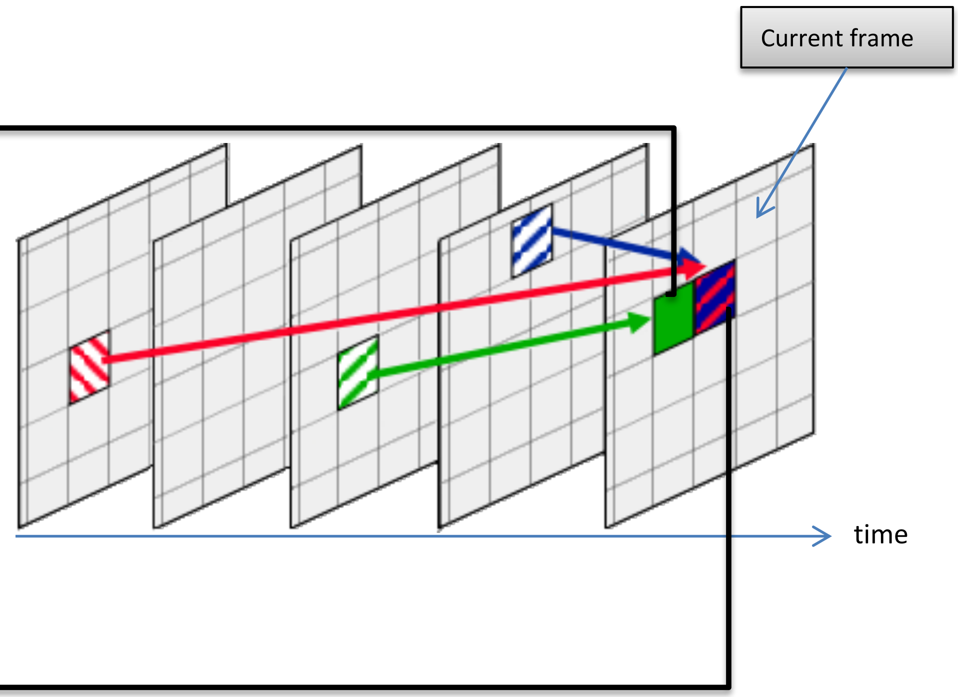
- Planar Prediction
  uses linear interpolation to generate the prediction signal



DC prediction

Inter-picture prediction

- Uni-Prediction
  - (P-Slice and B-Slice)
  - 1 reference picture

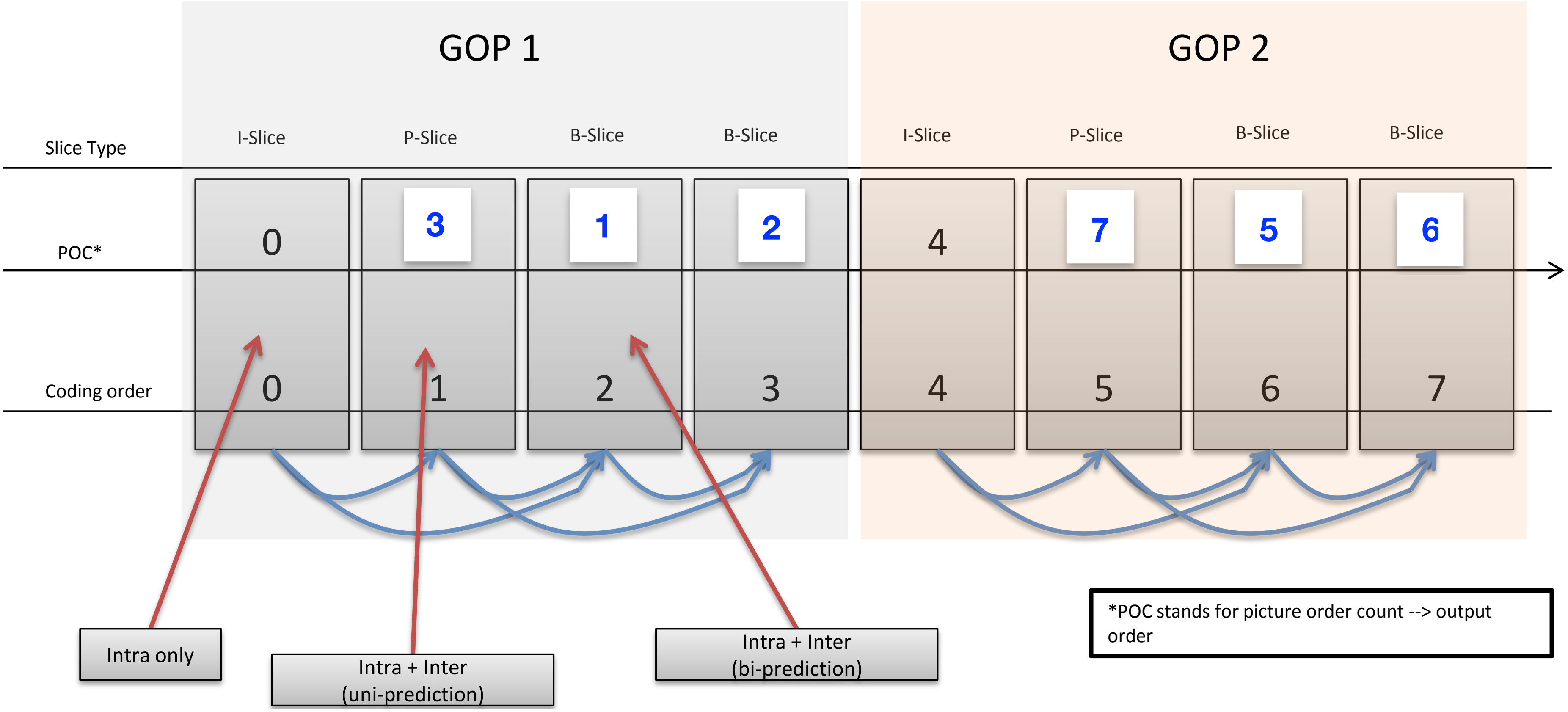


- Bi-Prediction
  - (B-Slice only)
  - 2 reference pictures

RPL: List of Reference Pictures for the current frame

Group of Picture (GOP) Concept

Specifies how the sequence of frames are encoded (Inter/intra)



HEVC advance motion vector prediction

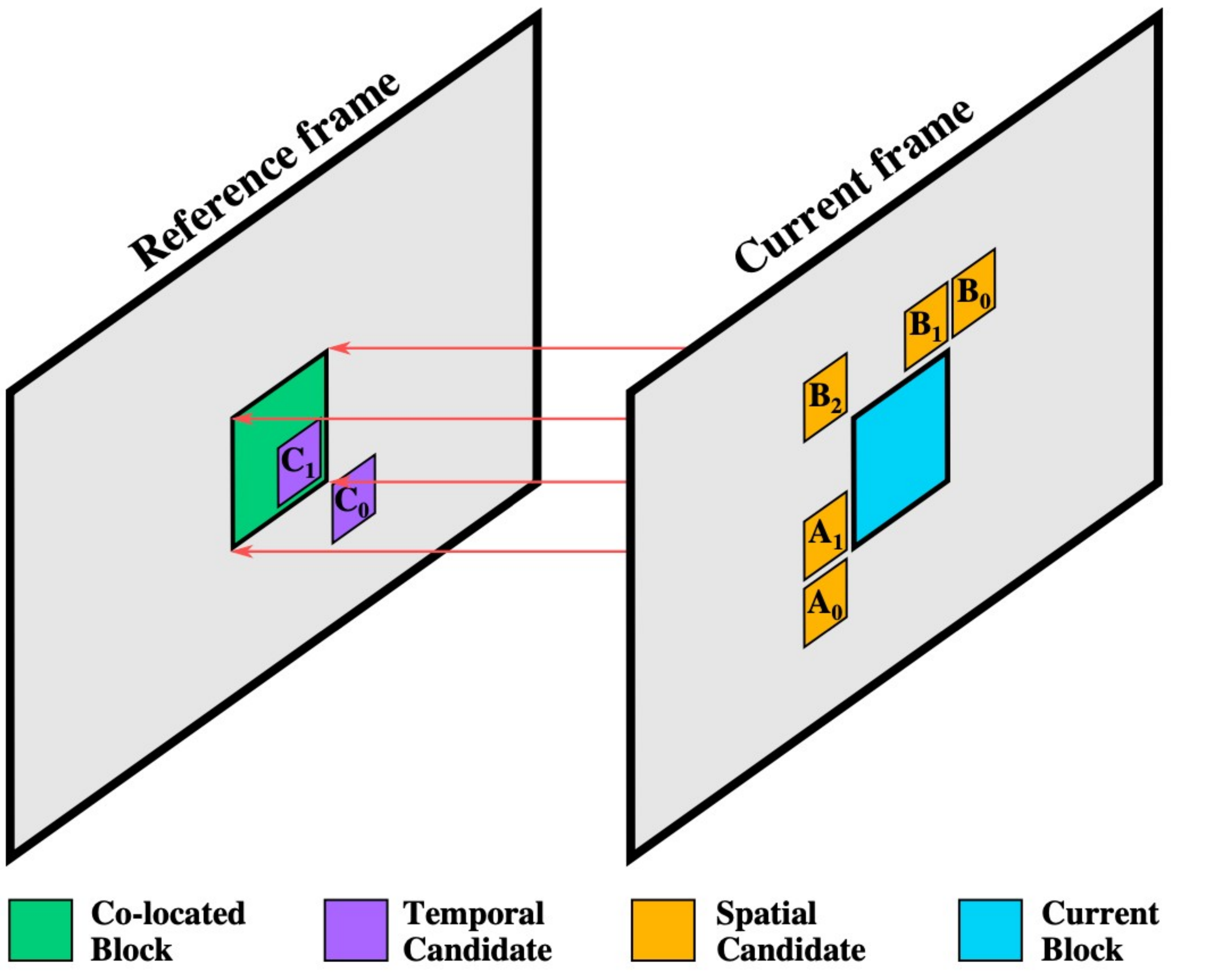


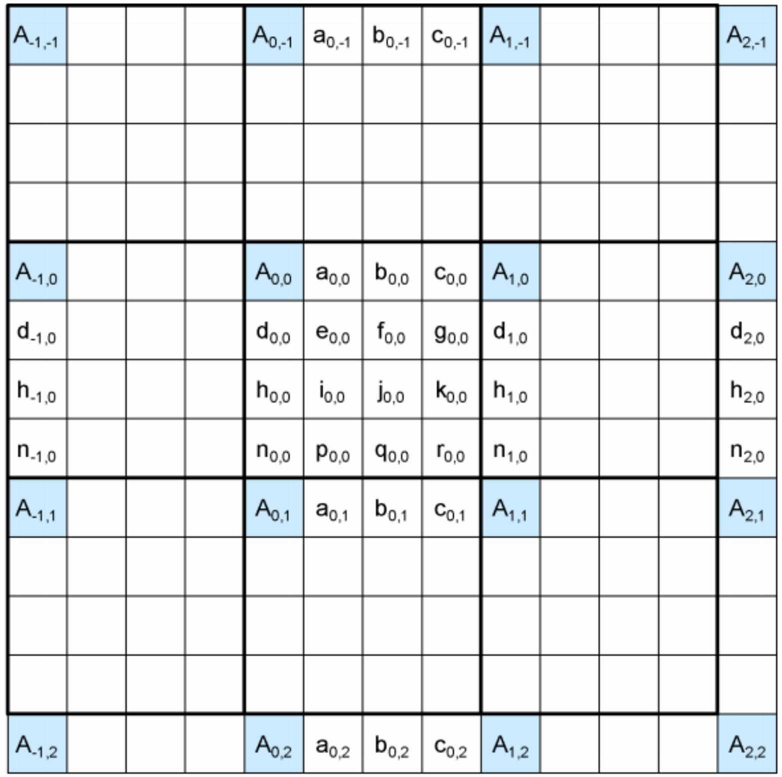
Fig. 2.12 Advanced Motion Vector Prediction (AMVP) candidates

Merge and Skip Mode

- When the motion information equals to the candidate's, HEVC encodes only a *merge* flag

- When the residual after motion compensated is neglected, HEVC encodes *skip* flag

Motion Compensation



HEVC uses ¼ sample resolution for motion compensation

Sullivan, Gary J., et al. "Overview of the high efficiency video coding (HEVC) standard." *Circuits and Systems for Video Technology, IEEE Transactions on*22.12 (2012): 1649-1668.

HEVC fractional sample prediction

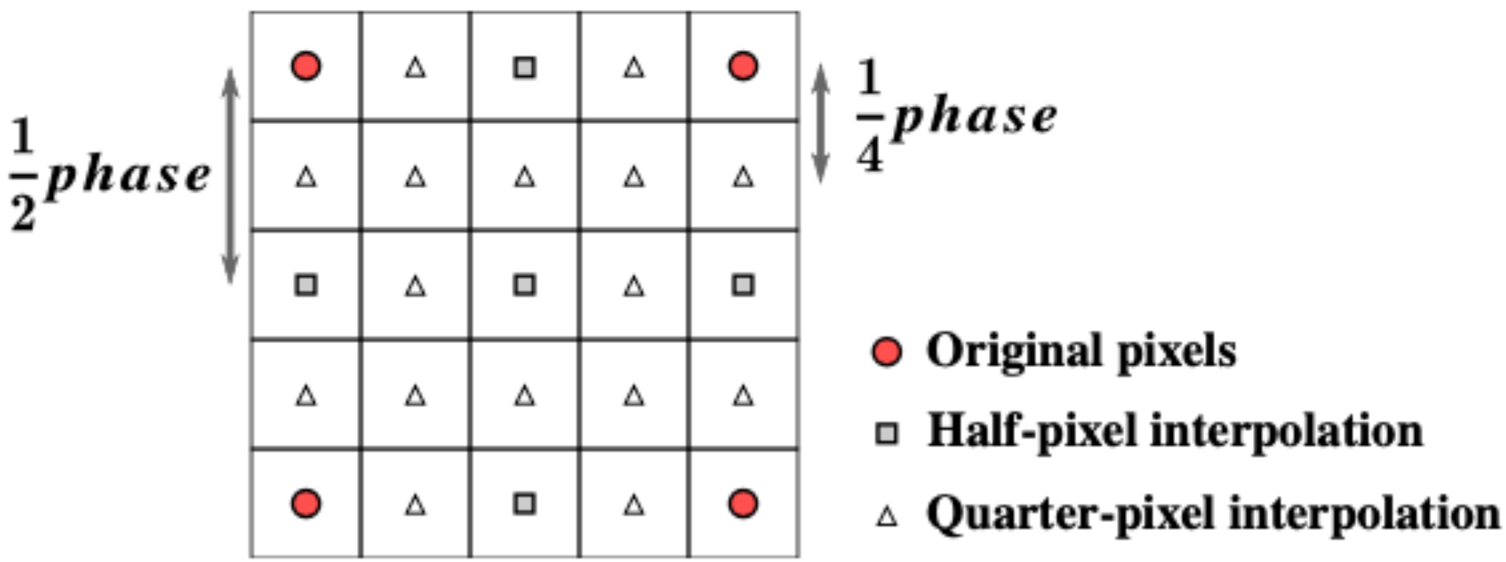
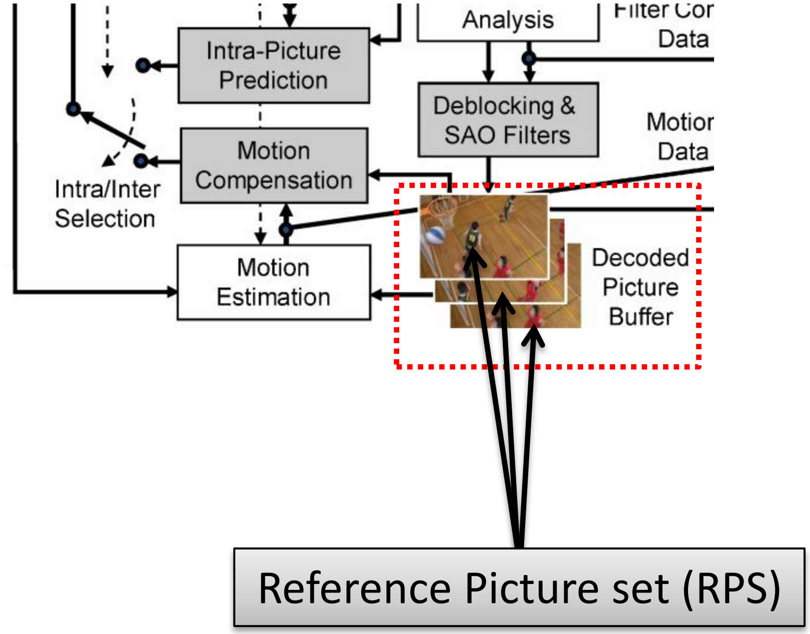


Fig. 2.11 Fractional sample positions for the interpolation of luma pixels.

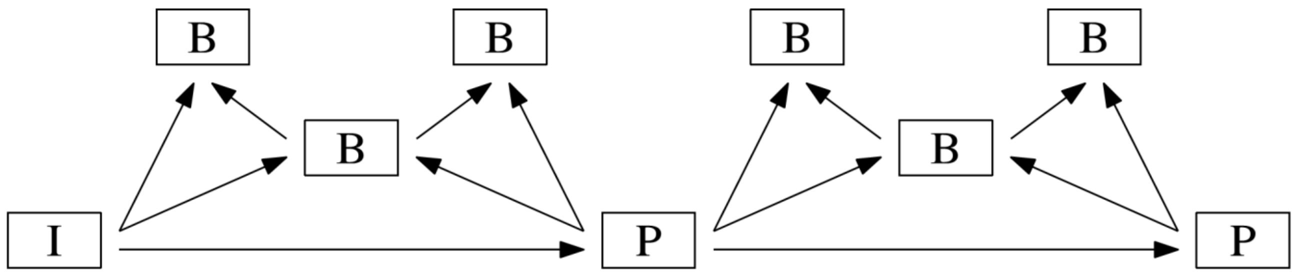
Decoded Picture Buffer

- Contains a set of previously decoded pictures, to be used for inter prediction
- Pictures in RPS that are used for inter prediction of the current image are listed in *reference picture lists*



Group of Picture (GOP) Concept

Ex: coding order != output order (From HM Doc.)

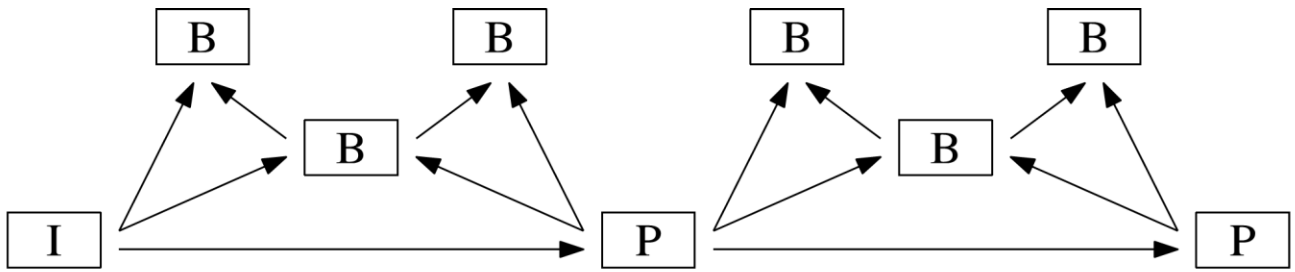


POC	0	1	2	3	4	5	6	7	8
Decode Order	0	3	2	4	1	7	6	8	5

POC: Picture Order Count = Display order

Decoding order = Coding order

Group of Picture (GOP) Concept



POC	0	1	2	3	4	5	6	7	8
Decode Order	0	3	2	4	1	7	6	8	5

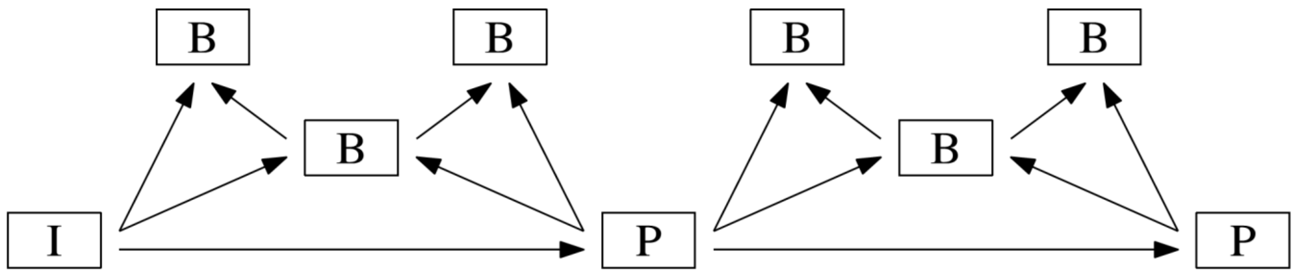
RPS RPL

Exercise: Write what is inside the reference picture set and reference picture list
(Decoded Picture Buffer size = 4)

RPS: set (series) of decoded pictures in the buffer (by using the POCs)
RPL: List of reference pictures (from the RPS) for a given frame

Group of Picture (GOP) Concept

Ex: coding order != output order (From HM Doc.)



POC	0	1	2	3	4	5	6	7	8
Decode Order	0	3	2	4	1	7	6	8	5
RPS :	XXXX	024X	04XX	0124	0XXX	3468	2348	4568	1234
RPL :	-	02	04	24	0	46	48	68	4

HEVC GOP prediction structure

TL: Temporal Layer

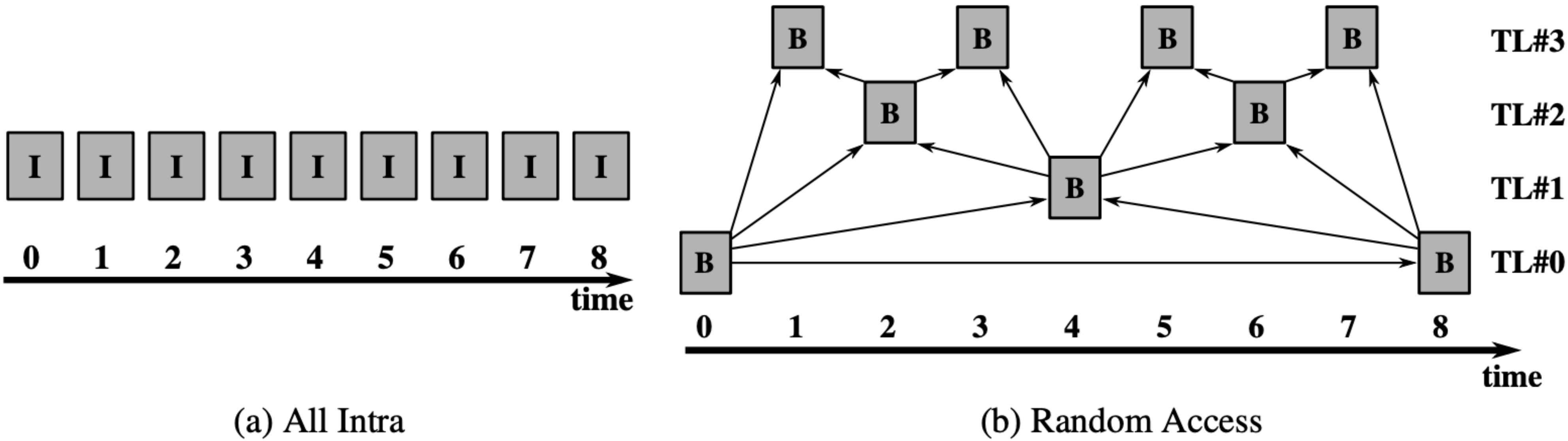


Fig. 2.15 GOP prediction structure for common HM configurations

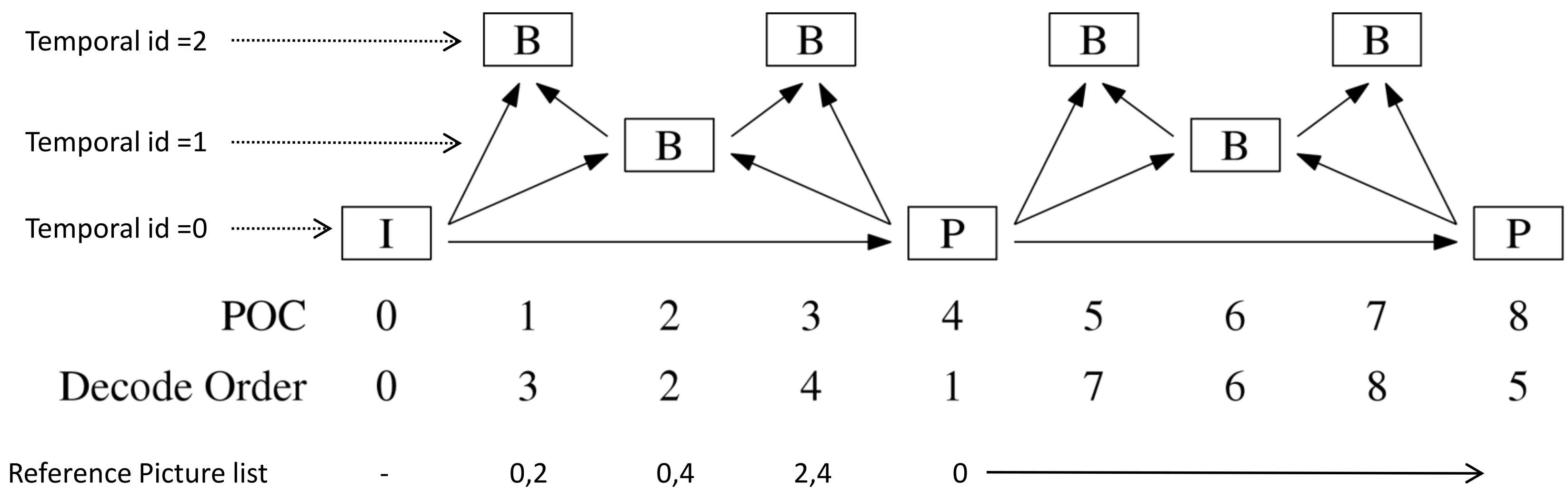
Temporal Layers

- Frames can be arranged in layers, identified by a temporal id.

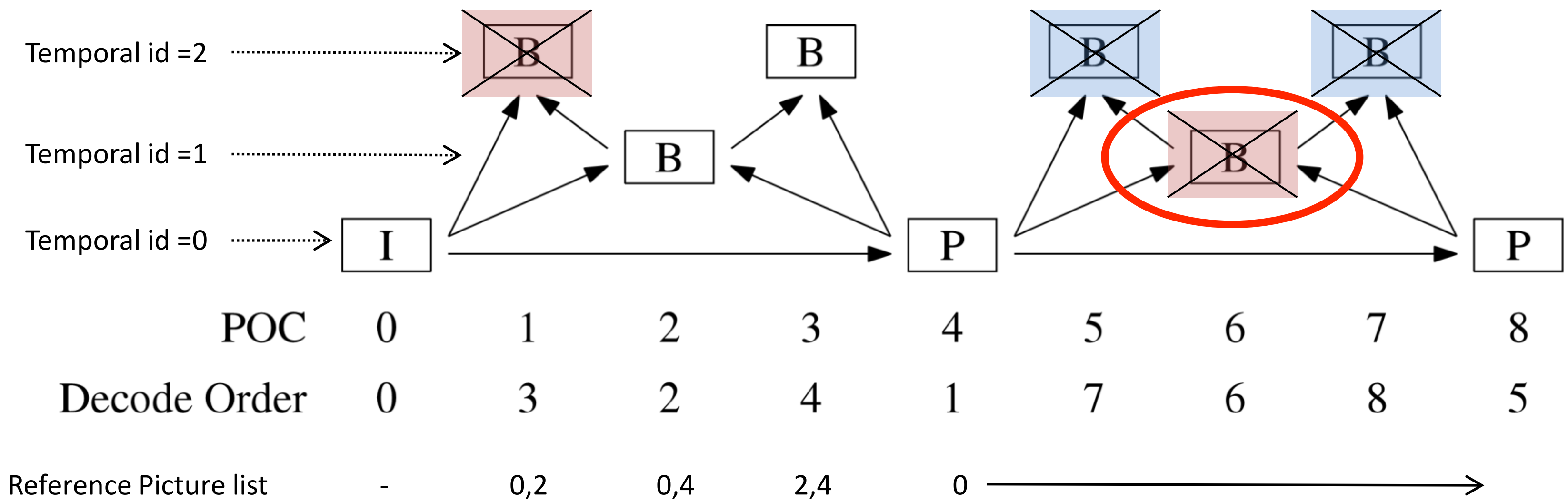
- Basic principle => Prediction from the lower layers

- Very useful in error concealment, as the errors in ~~higher~~ layers do not effect the ~~lower~~ layers.

Temporal Layers Example



Temporal Layers Example



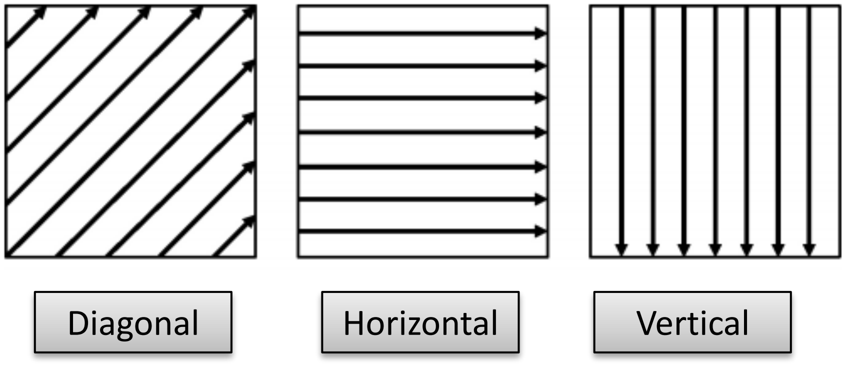
Lost Frames

Effected Frames

The video can still be decoded and viewed!

Transform in HEVC

- As mentioned, HEVC uses DCT (and DST for specific cases)



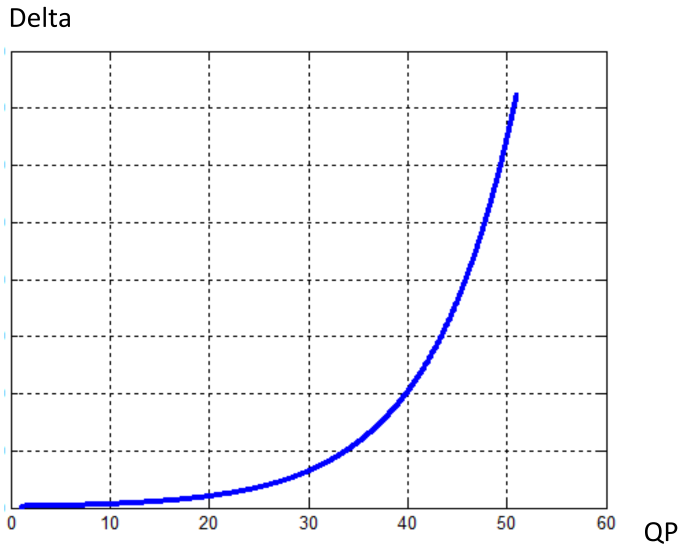
3 coefficients scanning available in HEVC

Quantization in HEVC

- Quantization step-size (Delta) is controlled by the *quantization parameter* (QP)

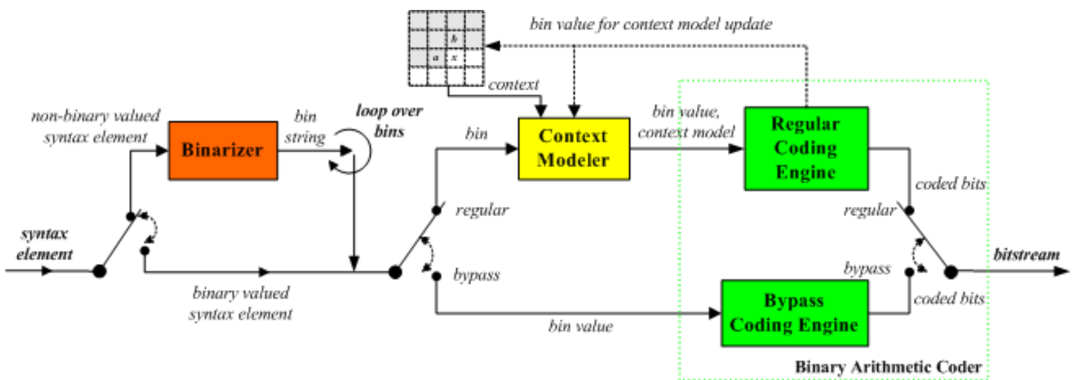
- Increasing QP by 6, doubles Delta

- QP range [0-51]



Entropy Coding in HEVC

- HEVC uses the Context Adaptive Binary Arithmetic Coding (CABAC)



- CABAC encodes the bit stream with number of bits as small as the average information (**entropy**)

Special Modes in HEVC

- PCM (pulse code modulation)

  Directly encode the samples, no transform, prediction & quantization.

  Useful when the signal characteristics are unusual where hybrid coding fails

Special Modes in HEVC

- Lossless mode

  if lossless compression is required, HEVC can switch off the lossy process (what is it?)

- How to achieve compression without losing information?

Special Modes in HEVC

- Transform Skip

  Transform can be skipped for a certain type of signals (eg. screen contents)

  Can only be applied to 4x4 TB

Reconstruction Filter

- Deblocking filter (DBP)

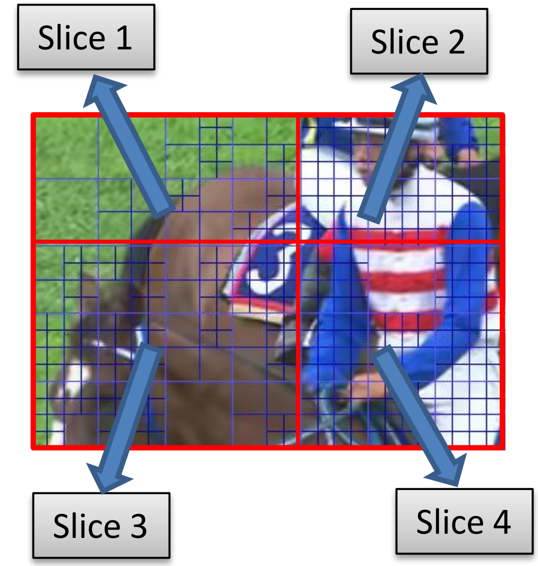
  Used to reduce blocking artifacts due to blocking process

- Sample Adaptive Offset filter (SAO)

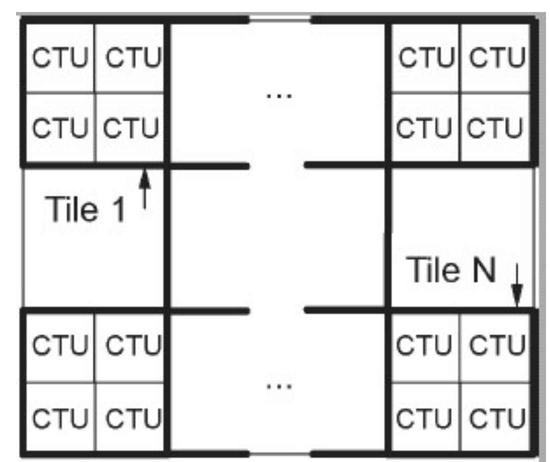
  Used to reconstruct lost edges based on neighboring information

Parallel Processing in HEVC

1. Slices: not in HM software

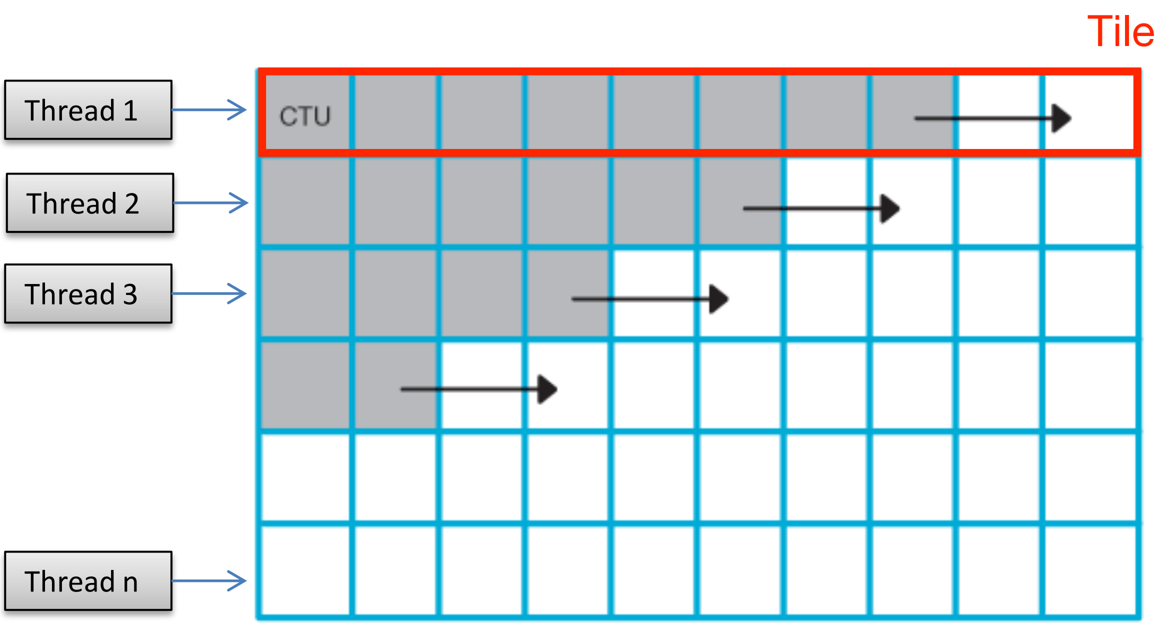


2. Tiles:
   – Groups of CTUs are arranged together such that no intra/inter prediction across tiles boundaries



Slice = 1 or more tiles

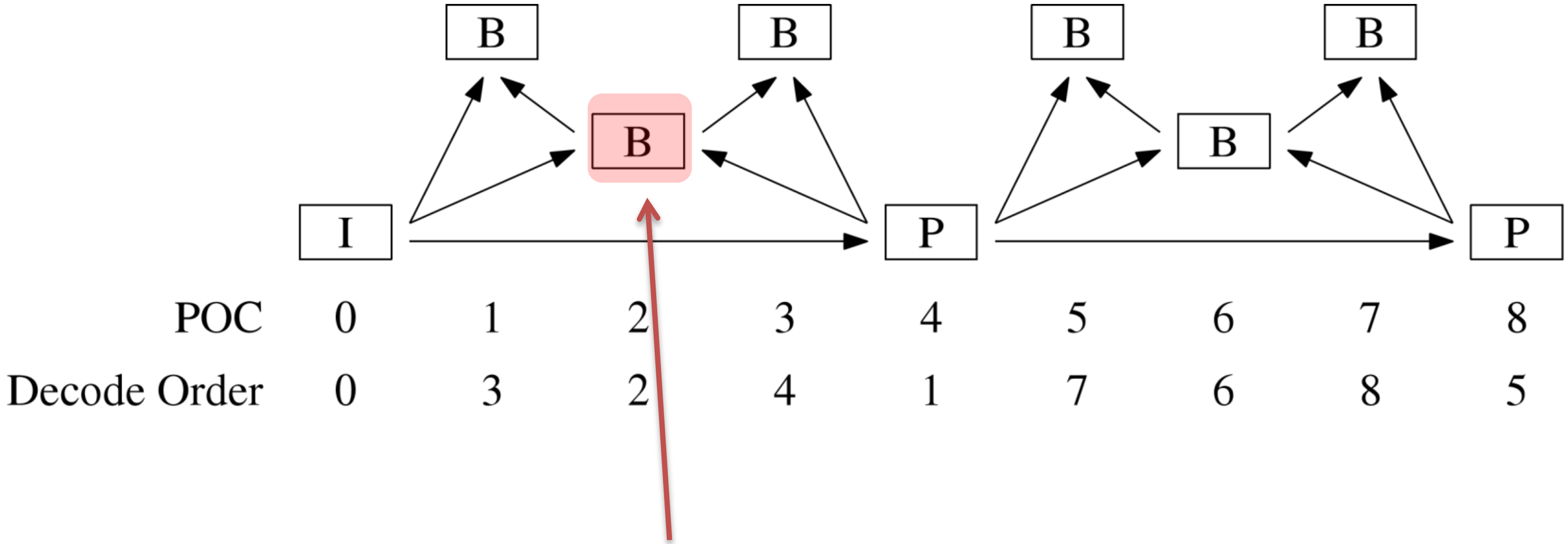
Parallel Processing in HEVC



Wave-Front Parallel Processing (WPP)

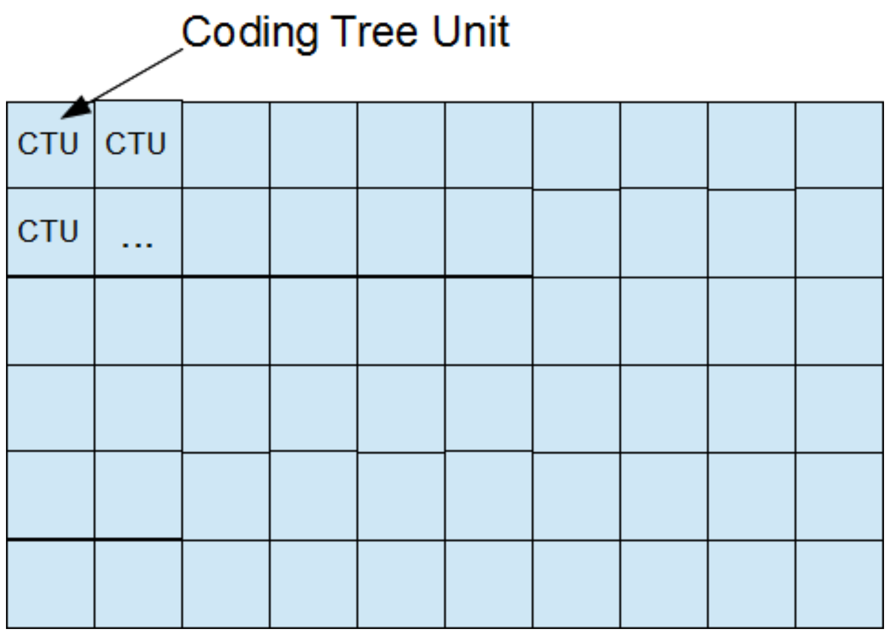
Example of Encoding Process

- Starting form GOP, lets pick one picture



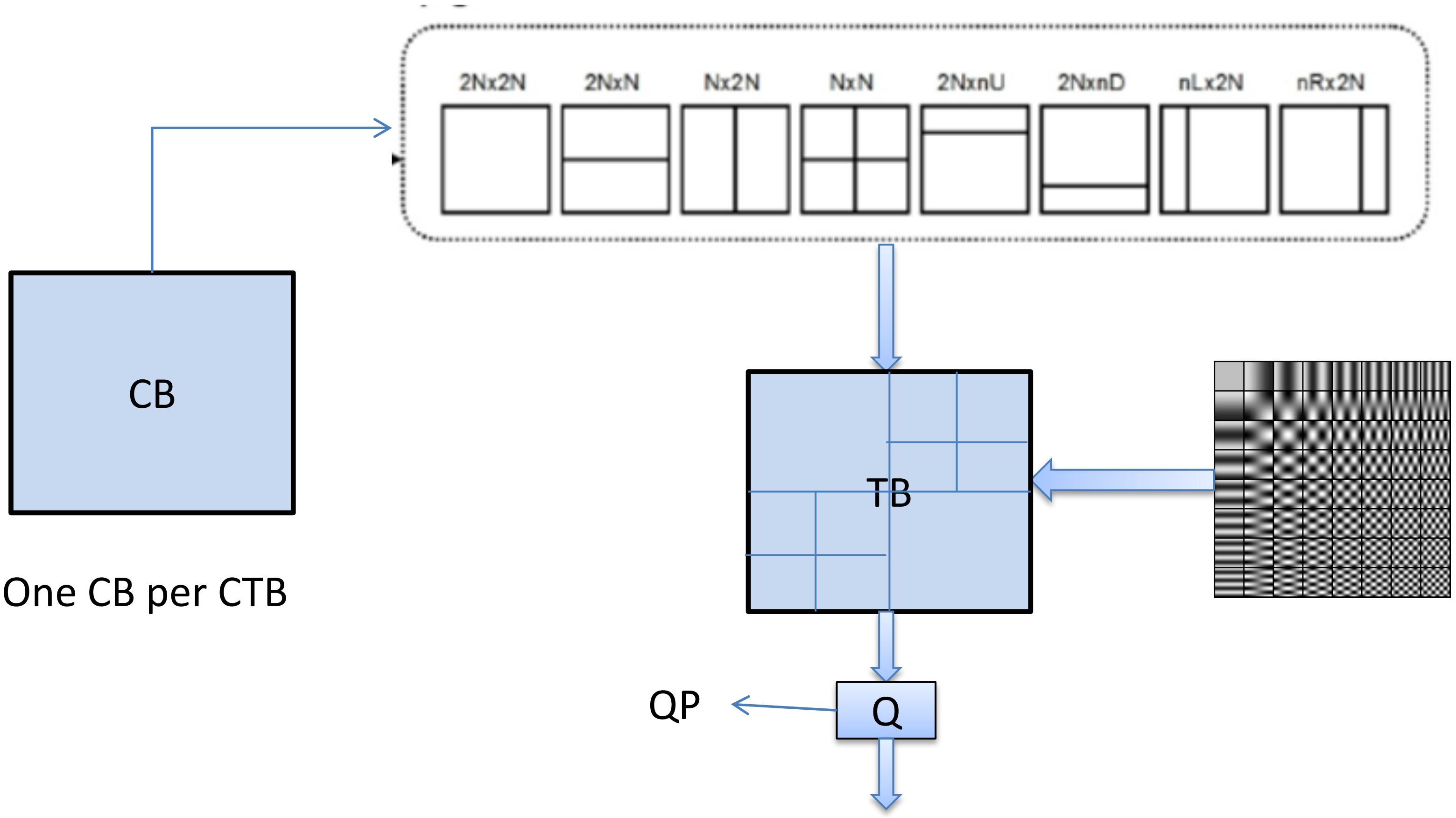
This is B-Picture:
- Intra and inter prediction are possible
- Uni and Bi directional are possible

Ex. Cont.



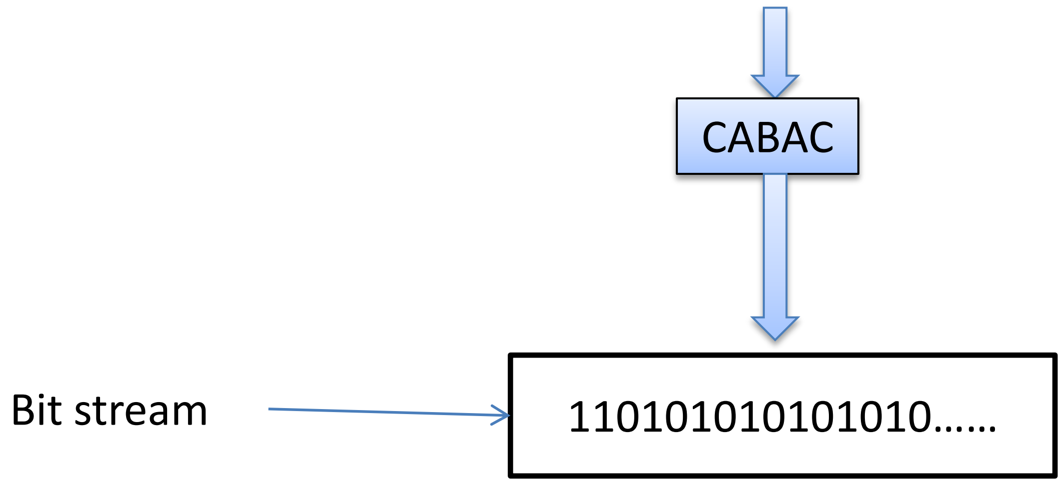
Divide the image into equal CTUs

Ex. Cont.



One CB per CTB

Ex. Cont.



Ex. Cont.

- CB can be also split, the process restart again at new depth (depth++)