Advanced Intra Coding
Intra-Picture Prediction

Transform coding: Typical design
- 2D Discrete Cosine Transform of type II (or integer approximation)
- Scalar quantization: URQs with same quantization step size $\Delta$
- Entropy coding (employing remaining statistical dependencies)
   - Can only utilize dependencies within transform blocks

Intra-picture prediction
- Can additionally utilize dependencies between transform blocks
- Very simple variant (JPEG and H.262 | MPEG-2 Video):
   - Predict DC coefficient using DC coefficient of previous block
- More advanced approaches can significantly increase coding efficiency

Two approaches of intra-picture prediction
- Prediction in transform domain
- Prediction in spatial domain (before transform coding)
Prediction of DC Coefficient

**Simplest Variant of Intra-Picture Prediction** (JPEG & MPEG-2 Video)
- Predict DC level with DC level of previous block (left to current block)
- Transmit only difference to predictor

\[ \Delta DC_k = DC_k - DC_{k-1} \]

- Reduces variance of pmf of transmitted DC level
- Improves coding efficiency

Example: Histogram with and without DC prediction (8×8 blocks)
Example: Effect of DC Prediction ($8 \times 8$ Blocks)
Advanced Intra Prediction in Transform Domain

Advanced Intra Coding mode of H.263: Three coding modes

- **DC prediction** and zig-zag scan
- **Horizontal prediction** and alternate-vertical scan
  - Suitable for blocks with mainly horizontal structures
- **Vertical prediction** and alternate-horizontal scan
  - Suitable for blocks with mainly vertical structures

Similar concept is also used in MPEG-4 Visual
How to Select Suitable Prediction Mode?

Mode Decision
- Syntax supports multiple intra prediction modes (mode is transmitted)
- Encoder has to select one of the supported modes
- Goal: Maximize coding efficiency

Lagrangian Mode Decision
- Evaluate all supported coding modes
  - Perform prediction, quantization, entropy coding, reconstruction
  - Calculate distortion $D = \sum_k (s_k' - s_k)^2$
  - Determine number of bits $R$ for transmitting mode and transform coefficient levels
- Select mode that minimizes Lagrangian cost function
  $$J = D + \lambda \cdot R$$
Example: Vertical prediction

- **Transform domain:** Predict first row of transform coefficients
- **Equivalent prediction in spatial domain**

\[
\hat{s}_{\text{ver}}[x, y] = \frac{1}{N} \sum_{k=0}^{N-1} s'[x, -1 - k]
\]

- **Simplified prediction in spatial domain**

\[
\hat{s}_{\text{ver}}[x, y] = s'[-1, y]
\]
Spatial Intra Prediction

Spatial intra prediction

- Similar complexity than similar operation in transform domain
- Usage of directly adjacent samples → Improved coding efficiency
- Main advantages:
  - Can also be applied if neighboring blocks are coded in an inter mode
  - Straightforward extension to multiple prediction directions
    (can include interpolation of border samples)

Intra prediction in video coding standards

- **H.262 | MPEG-2**: Predict DC coefficient from previous block
- **H.263 & MPEG-4**: DC, horizontal, vertical (in transform domain)
- **H.264 | AVC**: 9 spatial intra prediction modes (for $4 \times 4/8 \times 8$ blocks)
- **H.265 | HEVC**: 35 spatial intra prediction modes (for all block sizes)

→ Number of supported intra prediction modes is increased from one generation of video coding standards to the next
Spatial Intra Prediction in H.264 | MPEG-4 AVC
33 directional prediction modes (linear interpolation / ref. sample smoothing)

- DC prediction mode (similar to H.264 | MPEG-4 AVC)
- Planar prediction mode

⇒ In total: 35 spatial intra prediction modes
HEVC Planar Mode: General Idea

- Not all image blocks fit an edge model (well representable by directional image model)
- DC prediction only very coarse approximation (order-0 model)
- Alternative: Planar prediction mode

\[
\hat{s}[x, y] = \frac{N - 1 - x}{2N} \cdot s'[−1, y] + \frac{1 + x}{2N} \cdot \hat{s}[N - 1, y] + \\
\frac{N - 1 - y}{2N} \cdot s'[x, −1] + \frac{1 + y}{2N} \cdot \hat{s}[x, N - 1]
\]

\[
\hat{s}[N - 1, y] = \frac{N - 1 - y}{N} \cdot s'[N - 1, −1] + \frac{1 + y}{2N} \cdot \hat{s}[N - 1, N - 1]
\]

\[
\hat{s}[x, N - 1] = \frac{N - 1 - x}{N} \cdot s'[−1, N - 1] + \frac{1 + x}{2N} \cdot \hat{s}[N - 1, N - 1]
\]

\[
\hat{s}[N - 1, N - 1] = \frac{1}{2} \cdot s'[N - 1, −1] + \frac{1}{2} \cdot s'[-1, N - 1]
\]
HEVC Planar Mode: Actual Implementation

\[ \hat{s}_H[x, y] = (N - 1 - x) \cdot s'[−1, y] + (1 + x) \cdot s'[N - 1, -1] \]

\[ \hat{s}_V[x, y] = (N - 1 - y) \cdot s'[x, -1] + (1 + y) \cdot s'[-1, N - 1] \]

\[ \hat{s}[x, y] = \frac{1}{2N} \left( \hat{s}_H[x, y] + \hat{s}_V[x, y] \right) \]
Spatial Intra Prediction — Coding Efficiency

Experimental investigation with H.265 | MPEG-H HEVC
- Restricted to $8 \times 8$ blocks (effect of block size is discussed later)
- Limited number of used prediction modes (reference: DC prediction only)
  - Coding efficiency increases with number of supported intra prediction modes
Impact of block size selection for transform coding
- Coding efficiency of transform coding typically increases with block size
- Coding efficiency improvement becomes small beyond a certain block size
- Complexity increases with block size

Impact of block size selection for spatial prediction
- Correlation decreases with increasing sample distances
- Intra prediction is more effective for smaller block sizes
- Side information rate (for intra modes) increases with decreasing block size

Combination of intra prediction and transform coding
- Optimal block size depends on actual signal properties
- Natural images: Highly non-stationary statistical properties
  ➡️ No single optimal block size
  ➡️ Adaptive block size selection can improve coding efficiency
**Variable Block Sizes for Prediction and Transform Coding**

## Block Sizes in Video Coding Standards

**H.262 | MPEG-2 Video, H.263, MPEG-4 Visual**
- Fixed block sizes for prediction and transform coding
- $16 \times 16$ macroblocks (for signaling intra prediction mode)
- $8 \times 8$ transform blocks

**H.264 | MPEG-4 AVC (High profile)**
- $16 \times 16$ macroblocks
- 3 intra coding modes: Intra4x4, Intra8x8, Intra16x16
- Block sizes for prediction and transform coding: $4 \times 4$, $8 \times 8$, $16 \times 16$
- Intra prediction mode selected on basis of transform blocks
- Intra16x16: Only 4 prediction modes & low-complexity $16 \times 16$ transform
Picture partitioning into **coding tree blocks**

- Coding tree blocks (CTBs): Fixed size of $16 \times 16$, $32 \times 32$ or $64 \times 64$ luma samples
- Size of CTBs chosen by encoder
- Luma and chroma CTBs together with syntax are called **coding tree unit** (CTU)

Partitioning of coding tree blocks

- Quad-tree partitioning into **coding blocks** (CBs)
- Luma and chroma CBs together with syntax are called **coding unit** (CU)
- Minimum CU size: Selected by encoder, but equal to or larger than $8 \times 8$ luma samples
- Coding mode (intra or inter) is chosen for CU
- Coding order: Z-scan
Example: Picture Partitioning into Coding Units

Example for picture partitioning into coding units

- Picture with $2560 \times 1600$ luma samples of HEVC test sequence “Traffic”
- Quadtree-based partitioning into coding unit represents a simple scheme for locally adapting the block sizes to the image structure
Partitioning of a CB into **transform blocks** (TBs)
- Nested quad-tree partitioning
- TB corresponds to a single block transform
- Supported sizes: $4 \times 4$, $8 \times 8$, $16 \times 16$, $32 \times 32$
- Luma and chroma TBs together with syntax form a **transform unit** (TU)

Intra prediction and mode signaling
- One or four luma intra prediction modes per coding unit
- One chroma prediction mode per CU
- Actual intra prediction is performed transform block by transform block
- Improved prediction accuracy
Block Sizes for Intra-Picture Coding — Coding Efficiency

First coding experiment with H.265 | MPEG-H HEVC

- Reduce impact of intra prediction: Only DC prediction is enabled
- Check different fixed block sizes & variable block sizes
  - Fixed block sizes: Coding efficiency increases with block size
  - Variable block sizes provide coding gains
Variable Block Sizes for Prediction and Transform Coding

Block Sizes for Intra-Picture Coding — Coding Efficiency

Second coding experiment with H.265 | MPEG-H HEVC

- All intra prediction modes are enabled
- Prediction increases effectiveness of smaller block sizes
- Fixed block sizes: Medium block sizes provide best coding efficiency
- Variable block sizes provide coding gains
Block Sizes for Intra-Picture Coding — Coding Efficiency

Third coding experiment with H.265 | MPEG-H HEVC

- All intra prediction modes are enabled
- Start with $8 \times 8$ blocks and successively enable additional block sizes
  - Additional block sizes provide coding efficiency improvements
  - Beside intra-picture prediction, the support of additional block sizes is a main factor for the improvement in intra-picture coding
Part Summary

Transform coding of sample blocks
- Separable orthogonal transform: DCT or integer approximation
- Scalar quantization: URQs with same quantization step size $\Delta$
- Entropy coding: Utilize remaining dependencies between quantization indexes

Intra-picture prediction
- Utilize dependencies between transform blocks
- Two methods: Prediction in transform domain or spatial domain
- Spatial prediction: Straightforward realization of multiple prediction modes
- Coding efficiency typically increases with number of supported intra modes

Block sizes for intra prediction and transform coding
- Determine efficiency of prediction and transform coding
- Non-stationary character of natural images $\Rightarrow$ Variable block sizes
- Simple and flexible partitioning: Quadtree-based approaches
- Variable block sizes significantly increase coding efficiency