# Standardization in JVT: Scalable Video Coding

**Jens-Rainer Ohm**

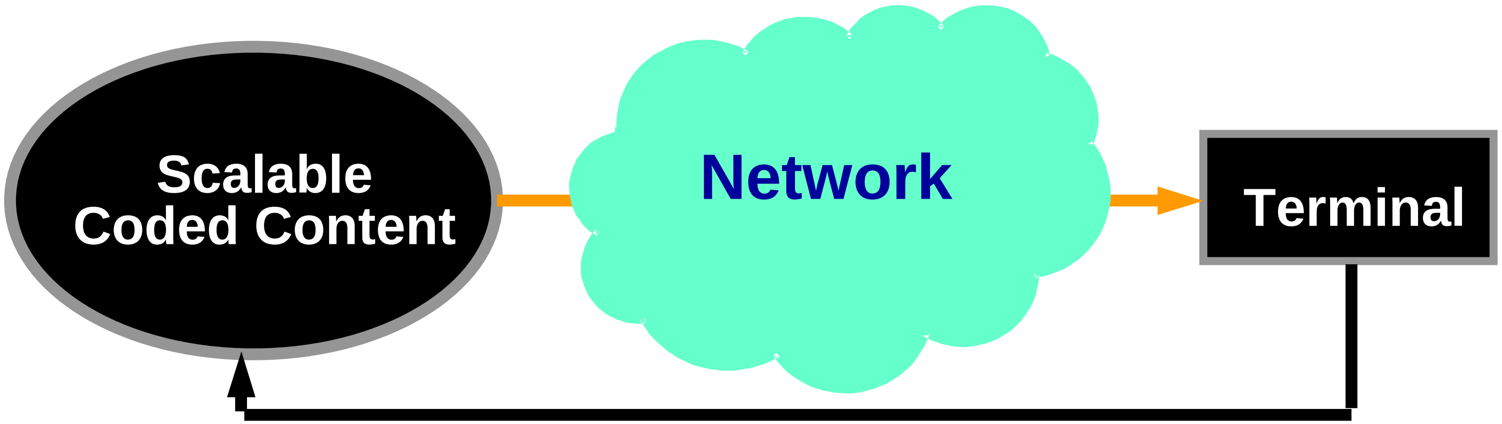
**RWTH Aachen University**
**Institute of Communications Engineering**
**ohm@ient.rwth-aachen.de**
**http://www.ient.rwth-aachen.de**

# Outline

- Introduction
- Summary on MPEG Call for Proposals on SVC and subsequent Core Experiments process
  - Results
  - Technical solutions
- Present status of SVC standard development in JVT
  - Temporal scalability and open-loop concepts
  - Spatial scalability
  - SNR and fine-granularity scalability
- Conclusions

# Scalable Media – the Idea

- **Universal Media Access:**
  code once and then customize the stream to access content
  - "Anytime"
  - from "Anywhere" (i.e. using any access network - wireless, internet etc.)
  - and by "Anyone" (i.e. with any terminal complexity)
- **Compatibility** of different formats/resolutions




Terminal capabilities & Network characteristics feedback

# MPEG/ITU-T Work Item on Scalable Video Coding

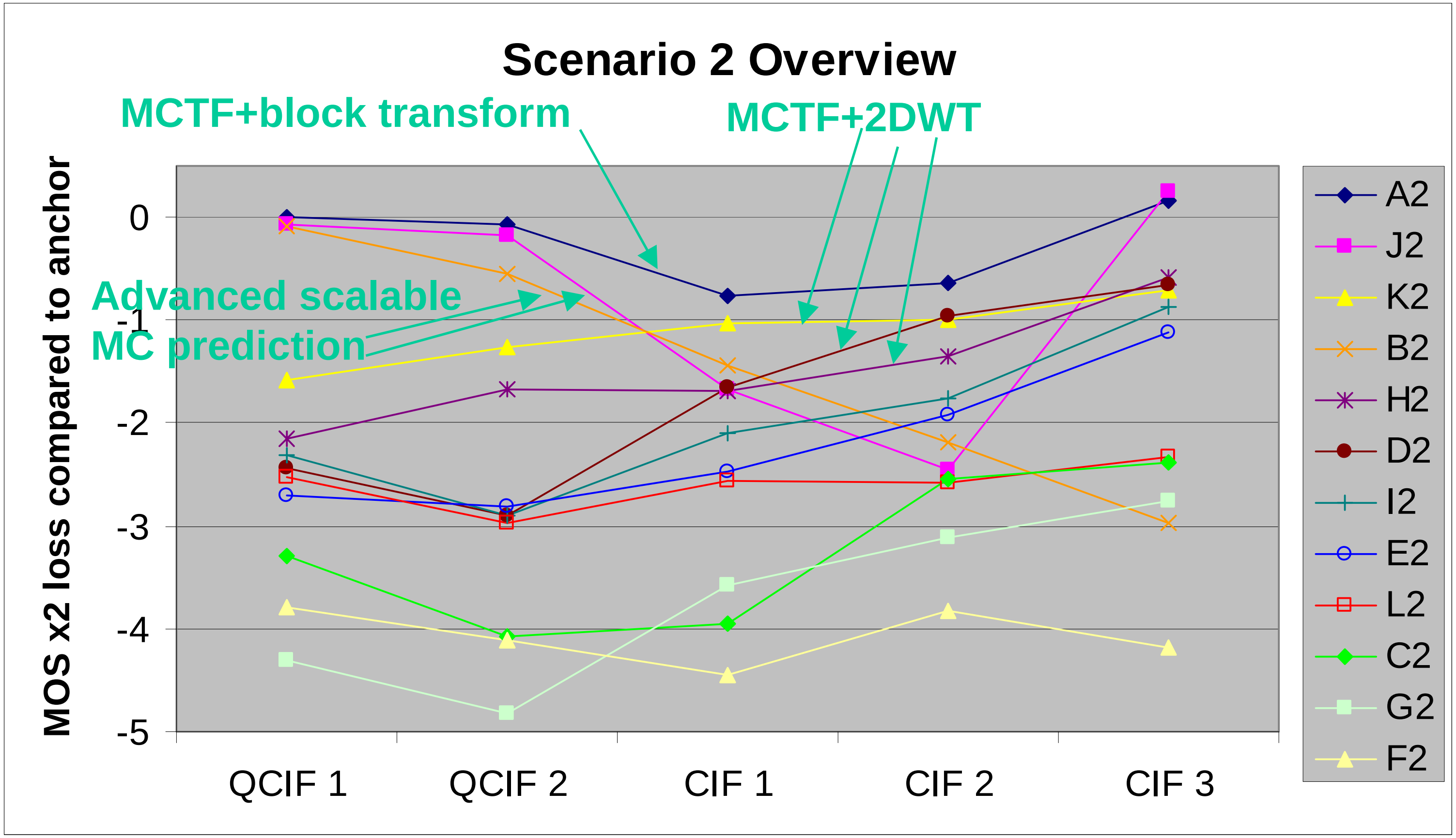
- SVC Work item
  - started as ISO/IEC 21000-13 (MPEG-21 Scalable Video Coding)
  - was moved to ISO/IEC 14496-10 January 2005
- Timeline:
  - October 2003: Call for Proposals
  - March 2004: Evaluation of proposals
  - January 2005: Working Draft (now within JVT)
  - October 2005: Committee Draft
  - March 2006 Final Committee Draft
  - July 2006 Final Draft International Standard



JVT

# MPEG Call for Proposals on SVC – Results

- Scenario 2: Scalability over 2 spatial layers




JVT

## Following the Call

- Core experiments were run to further identify and develop SVC technology
  - based on Scalable Video Model (SVM)
- Comparison made for wavelet-based and AVC/H.264-extending solutions
  - Full range of scalabilities tested (3x spatial, 3x temporal, medium-granularity SNR)
  - Formal visual tests (experts & non-experts) indicated superiority of AVC/H.264 solution, particularly at low rates and resolutions

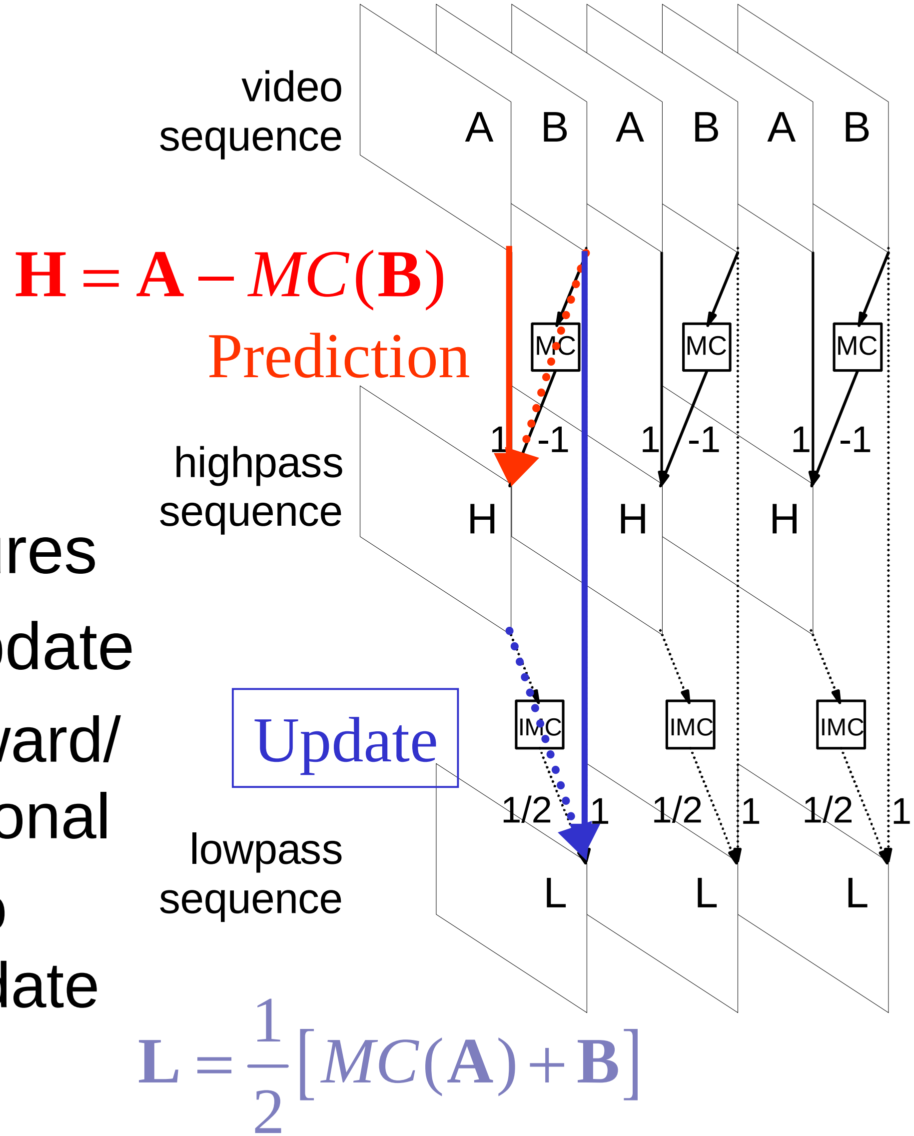
# JSVM and Working Draft

- JSVM: Joint Scalable Video Model (extending from MPEG Scalable Video Model 3.0)
- Working Draft of 14496-10:2005/AMD1 Scalable Video Coding
  - Extension of AVC / H.264 with compatible base layer
  - Motion-compensated temporal filtering (MCTF) with adaptive prediction and update steps for open-loop compression
  - Layered structure with "bottom-up" prediction from lower layers
  - One decoder loop
  - FGS functionality as extension of CABAC, modified interleaved scan order („cyclic block coding“)
  - New SVC-specific NAL unit types
  - Bitstream scalability at level of NAL packets

JVT

# Motion-compensated Lifting Filters

- Temporal-axis lifting filters, MC/IMC in lifting flow
  - MC and IMC must be aligned
- H pictures similar to MC prediction P pictures
- Adaptive predict & update
  - Switching intra/forward/ backward/bi-directional
  - L pictures similar to I pictures when update turned down

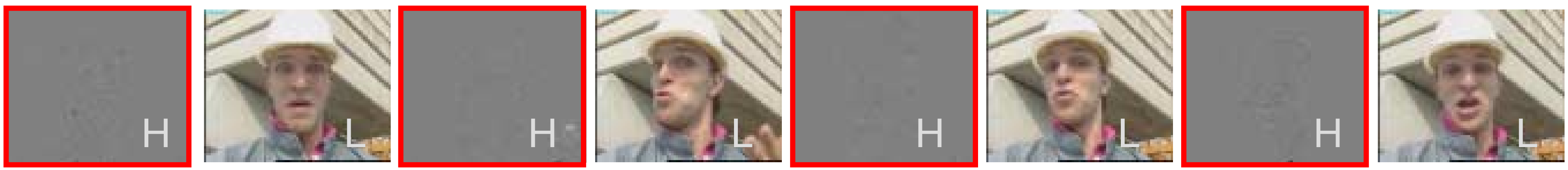




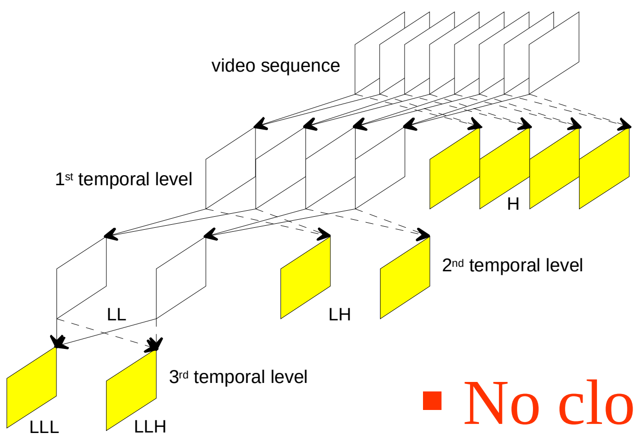

JVT

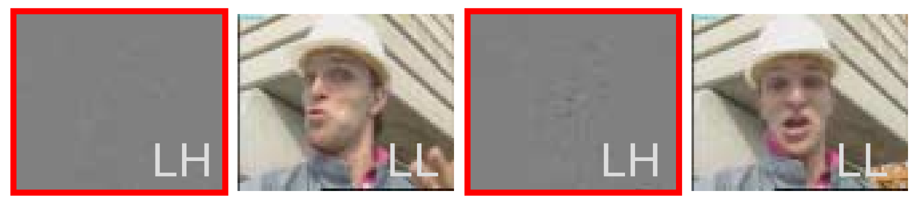
# MCTF as „Temporal-axis Wavelet Tree“


original video sequence




1st decomposition level



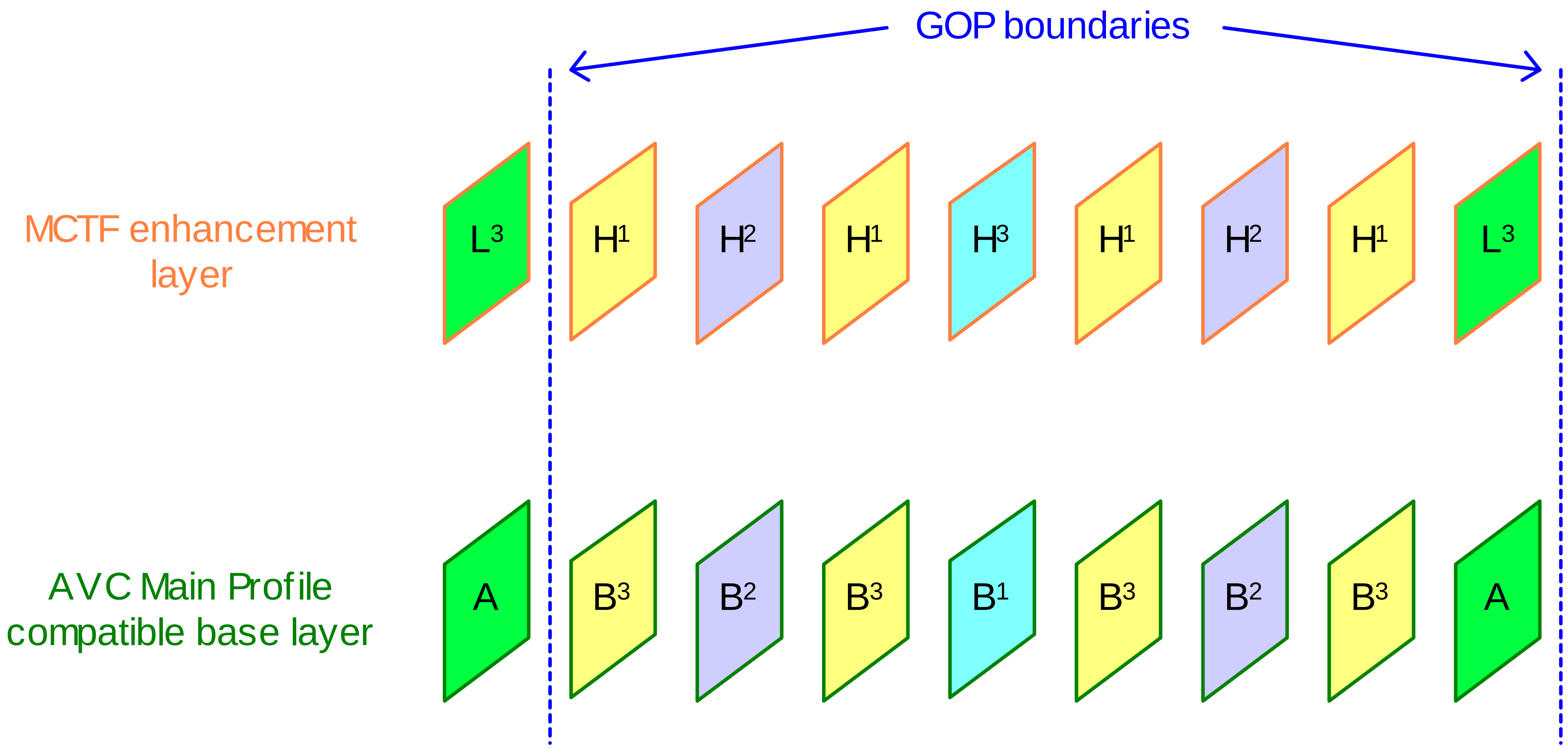

2nd decomposition level


3rd level

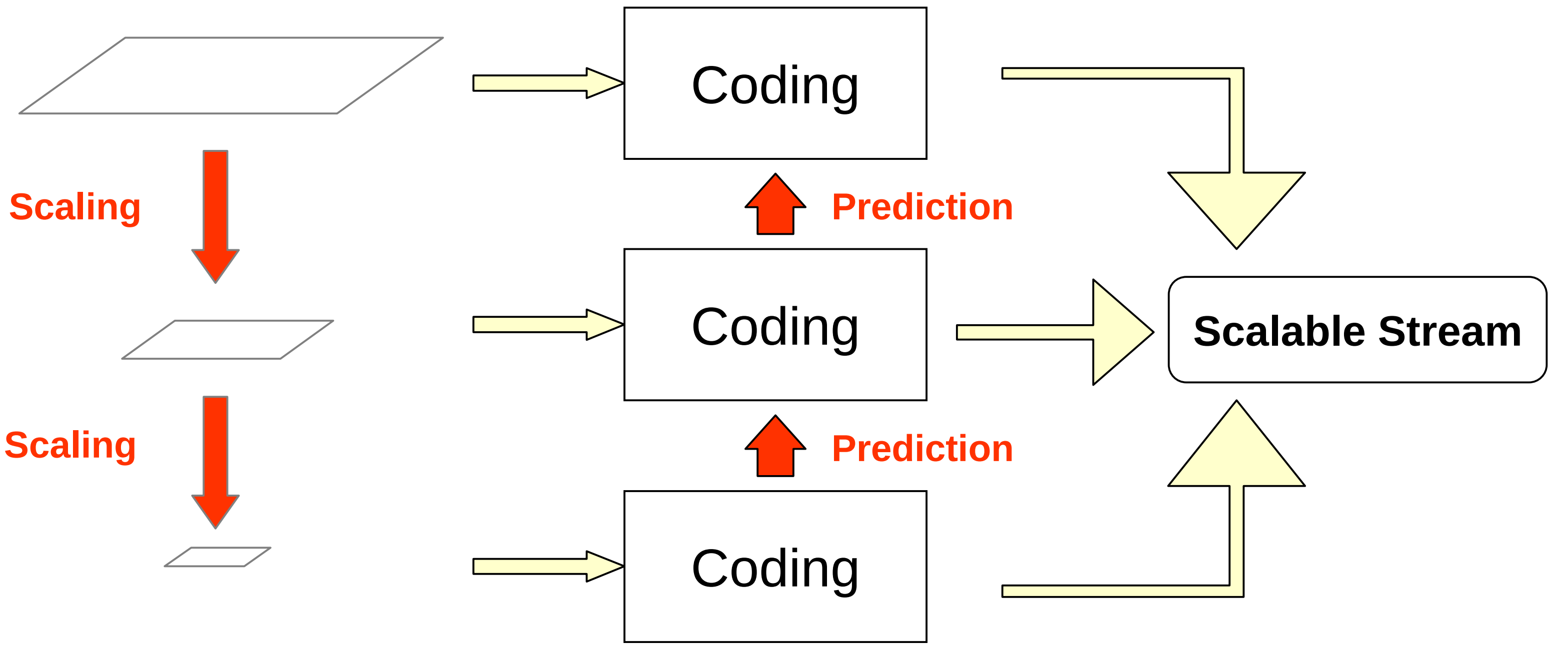
- No closed loop in prediction and update required
- Prediction loop over LLL.. possible

JVT

# Layered Coding with compatible base layer

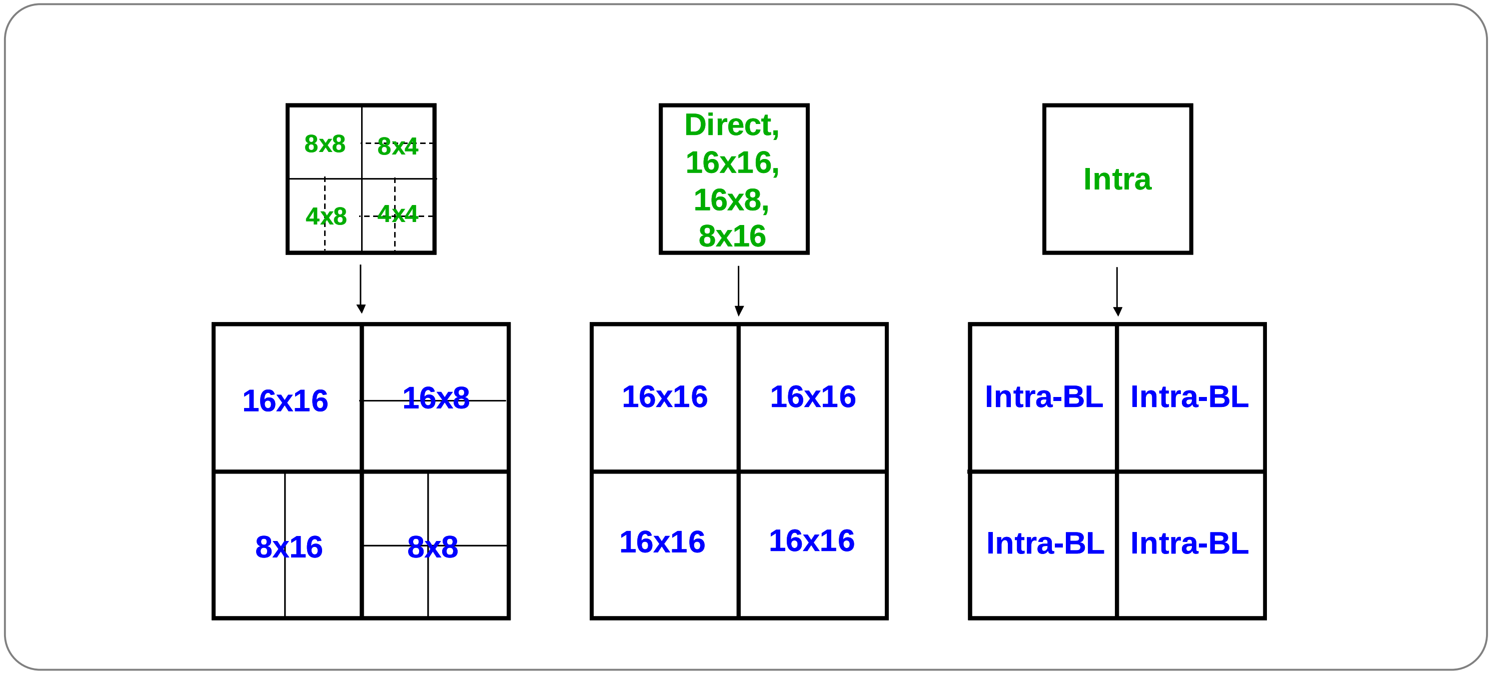


# Layered Coding with Spatial Pyramids




- **Pyramid structure**
  - Basic building blocks from AVC / H.264
  - Spatial transform coding using integer DCT approximations (4x4 and 8x8)

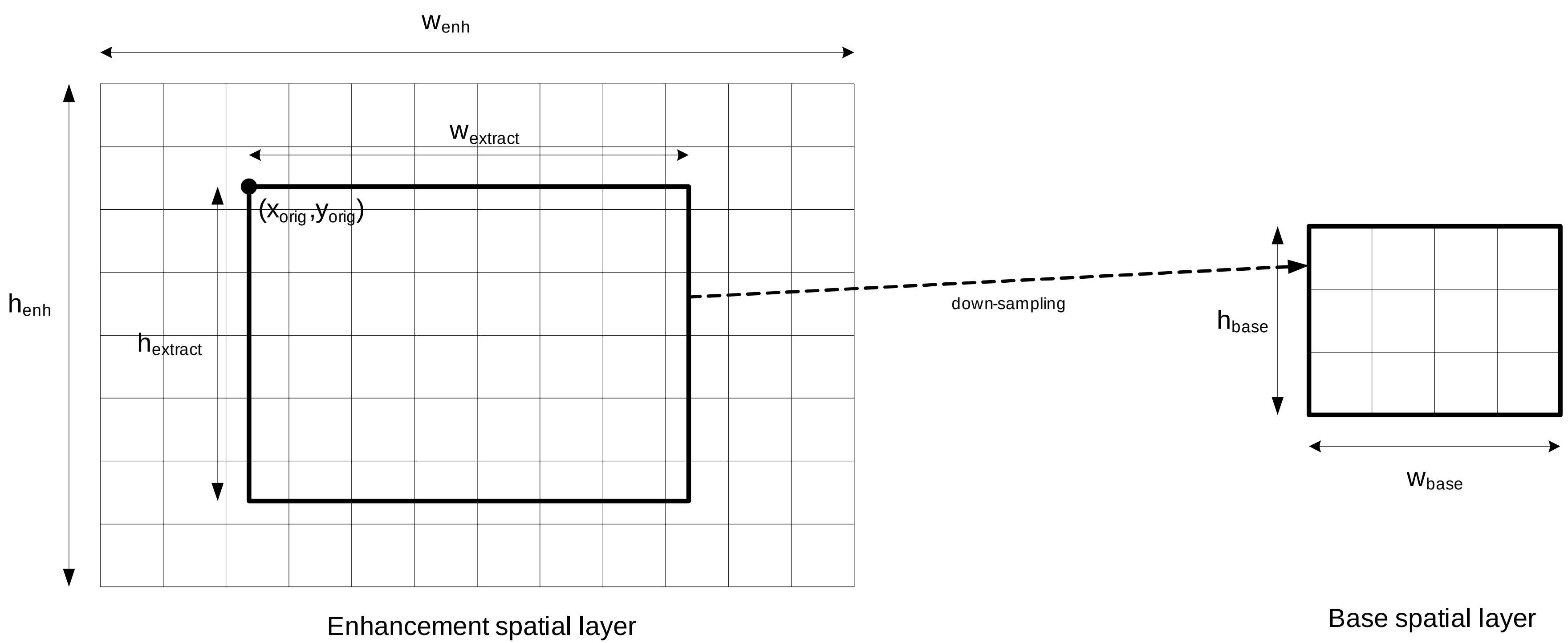
# Inter Layer Prediction




- **Motion:** Upsampled partitioning and motion vectors for prediction
- **Residual:** Block-wise upsampled residual (bi-linear)
- **Intra:** Upsampled intra MB (H.264 / AVC 1/2-pel filter)

# Extended Spatial scalability

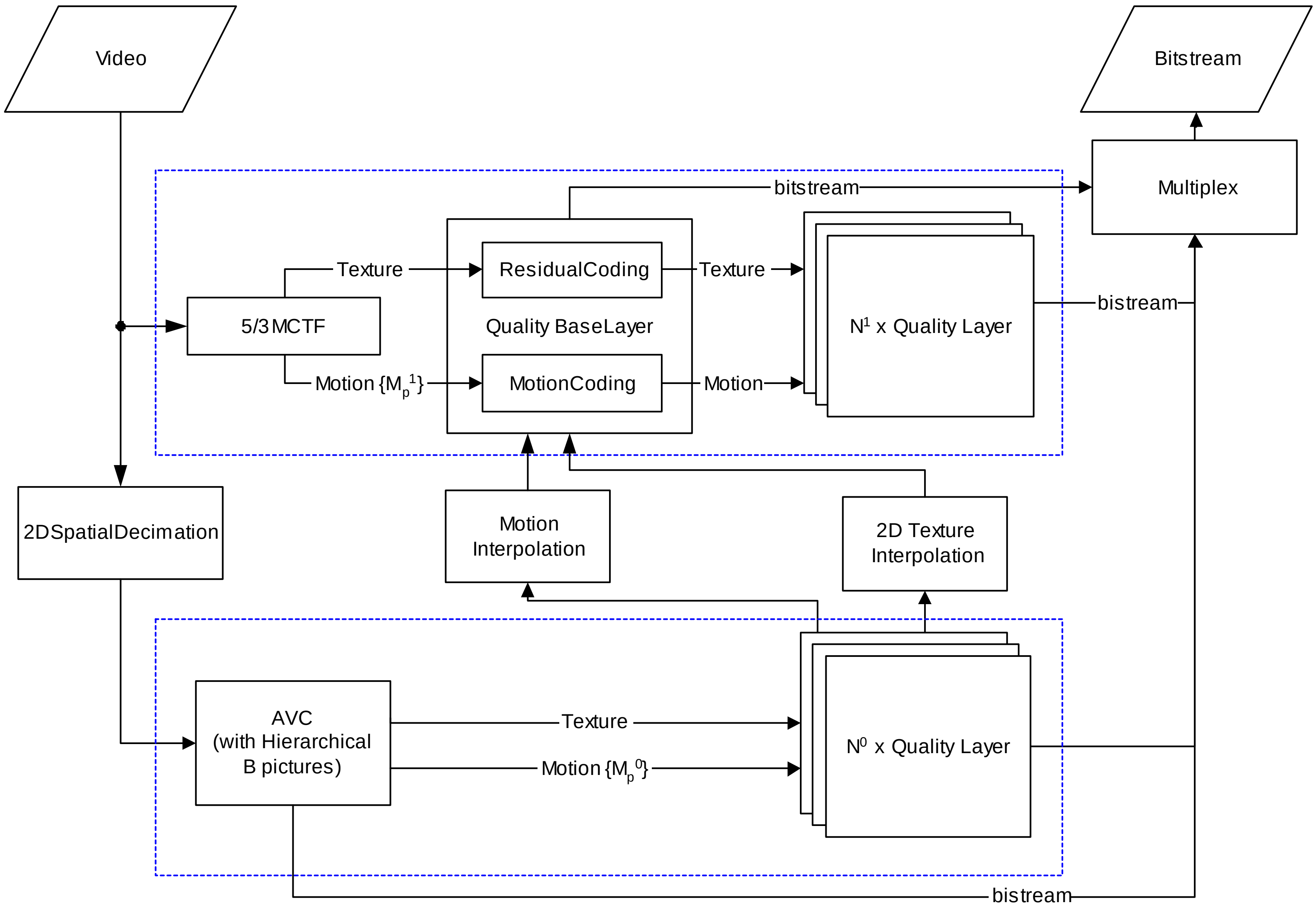
- Cropping and non-dyadic up-sampling
- Macroblocks not aligned between layers



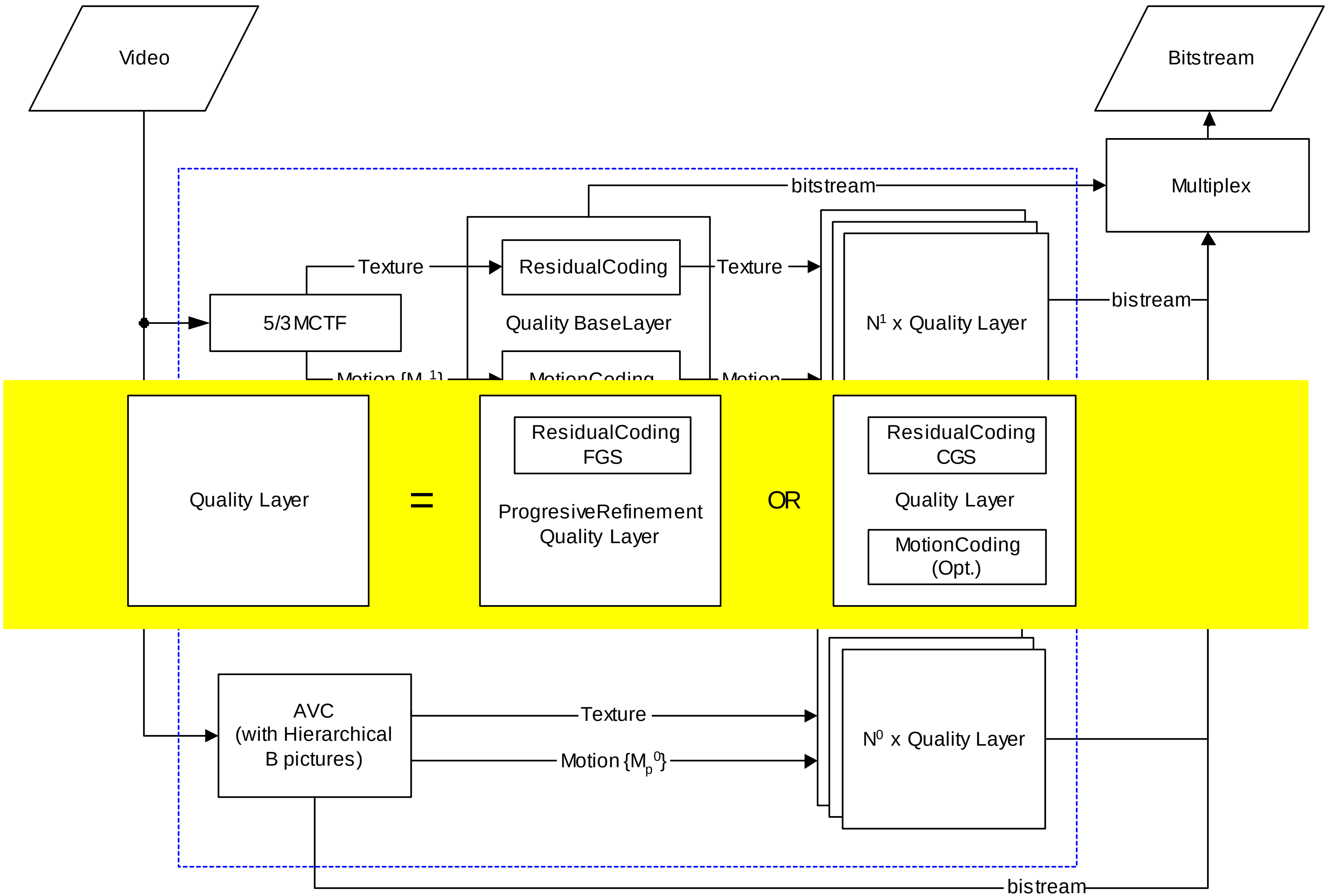
# SNR Scalability

- **Layered representation of L/H pictures**
- **Non-scalable 'base-layer'** (BL)
  - Motion information
  - Coarsest represenation of intra and residual data
  - Minimum acceptable reconstruction quality
- **Quality scalable enhancement layers** (EL)
  - Residue between original and SNR-BL representation
  - Doubled quantizer precision per SNR-EL
  - FGS: truncation of EL packets at arbitrary points
  - Refinement in the transform domain
  - Single inverse transform at the decoder side
- **"Dead substream" concept** allows flexible increase of quality beyond the limitations of layered coding

# SNR&Spatial scalability: 2 spatial layer example



# SNR&Spatial scalability: 2 spatial layer example

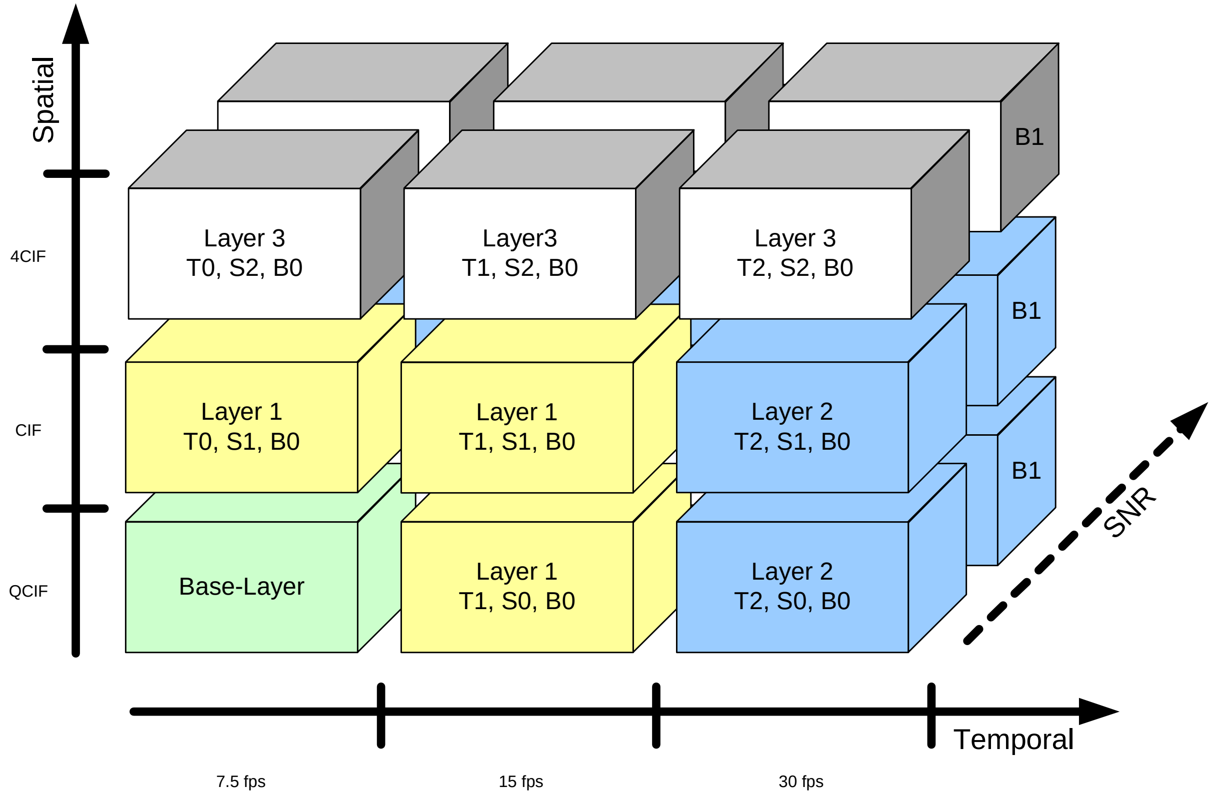


# Realisation of FGS SNR Scalability

- **Three scans for FGS coding of transform coefficients**
  1. Non-significant transform coefficients of significant blocks
  2. Refinement symbols for already significant coefficients
  3. Coefficients of non-significant blocks
- **Scanning pattern for each scan**
  - Scanning from low to high frequency bands (zig-zag)
  - Inside each band: first luma, then chroma in raster scan
- **Coding of transform coefficient levels**
  - Non-significant coefficients
    - Coded Block Pattern, Coded Block Bit, DeltaQP, Transform Size
    - CABAC symbols (SIG, LAST, SIGN, ABS)
  - Significant coefficients
    - Refinement symbol (-1, 0, +1)

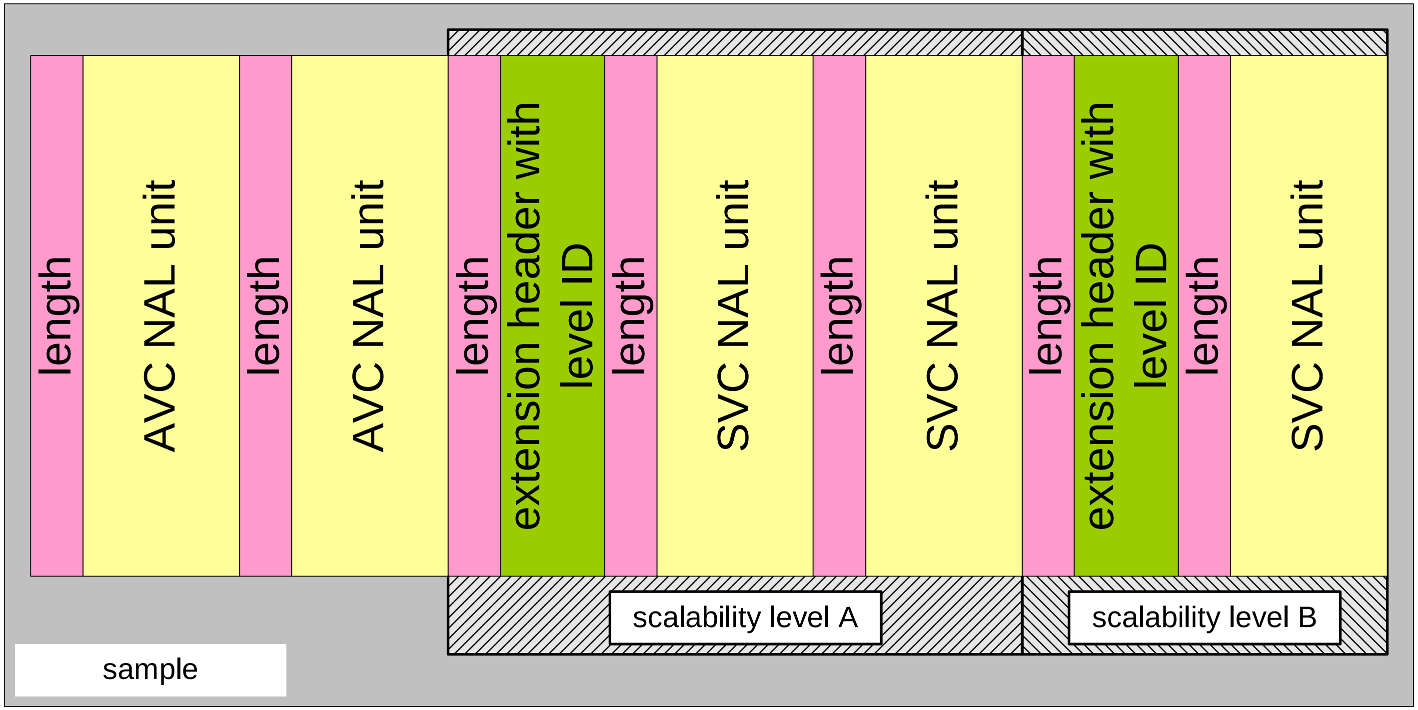
# Bitstream structure

- Data cube model and layered stream representation

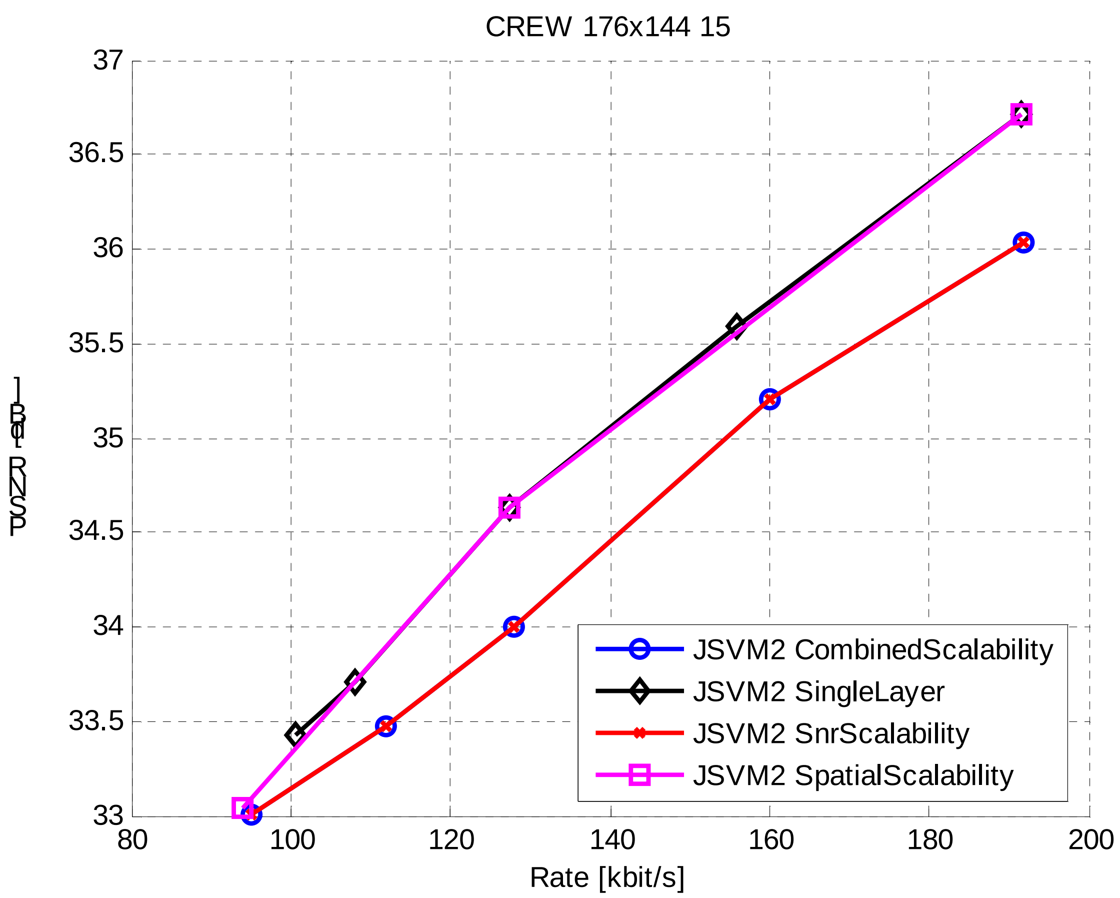


- New SVC NAL unit types can be ignored by conventional AVC parser

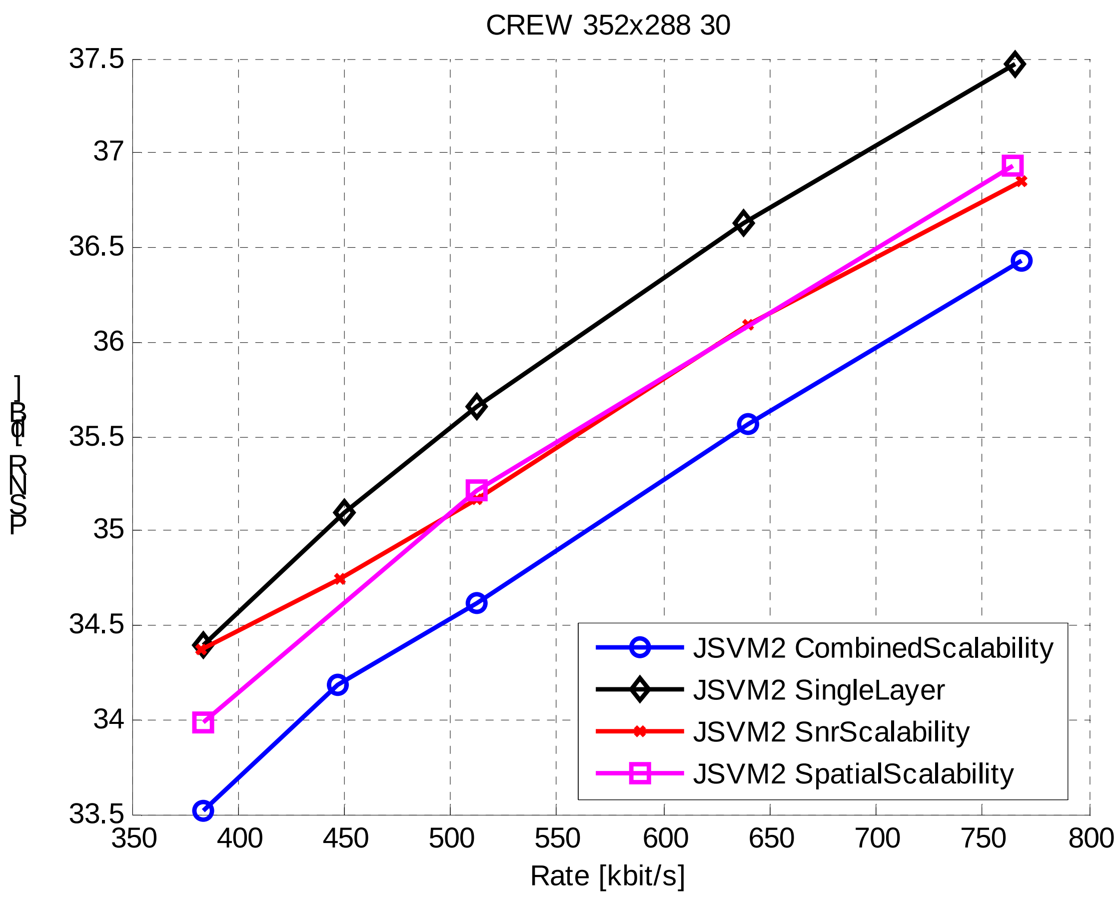

SVC sample



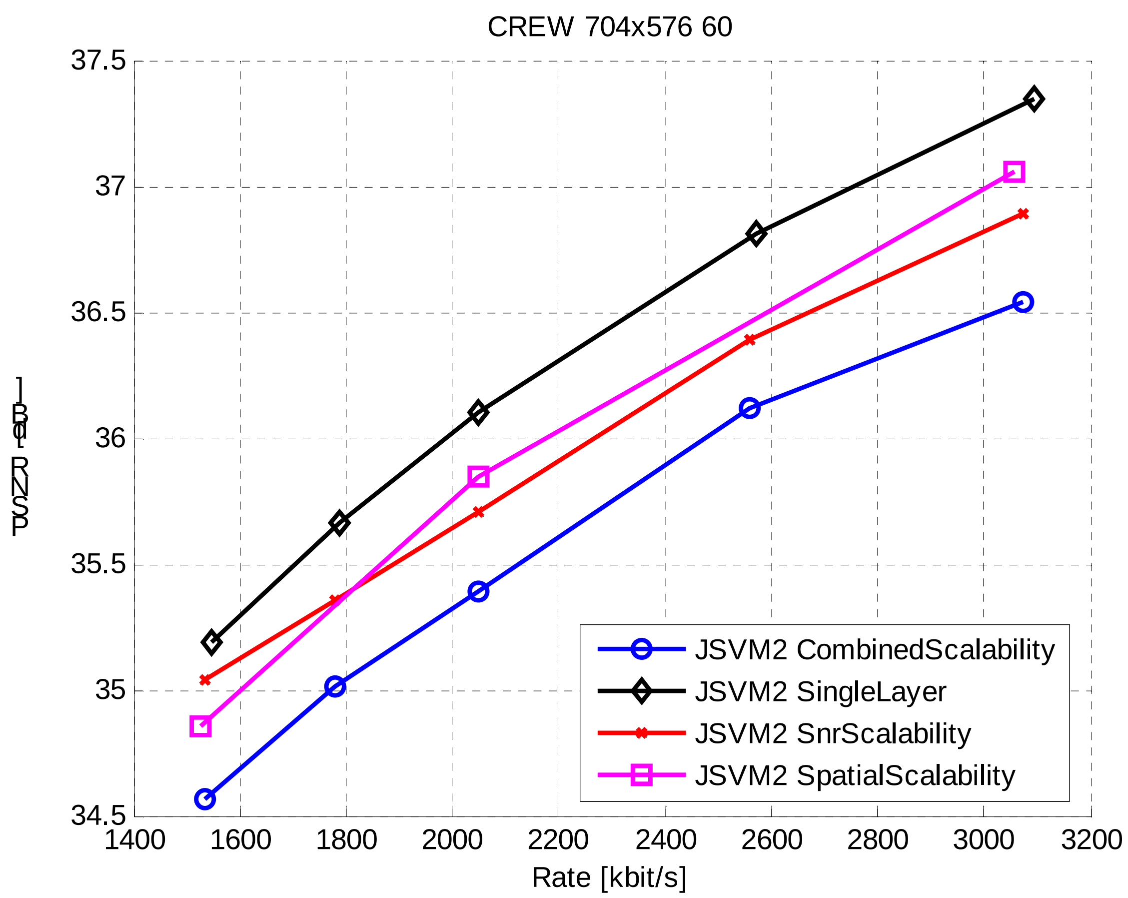
# Performance Example



# Performance Example



# Performance Example



# Summary

- **SVC to become an extension of H.264 / AVC**
  - Residual coding using existing tools
  - MCTF open loop structure with update step
  - Pyramid structure with inter layer prediction
  - FGS quality layers plus a minimum quality base layer
- **Good compression performance**
  - tradeoffs by the amount of flexibility in scalability
- **Further improvements expected e.g. by**
  - adaptive MCTF structures
  - combinations closed/open loop
  - improvement of quantization & FGS schemes
  - cross-layer dependencies of texture and motion
  - joint RD optimization over multiple layers