ECE462 – Lecture 28

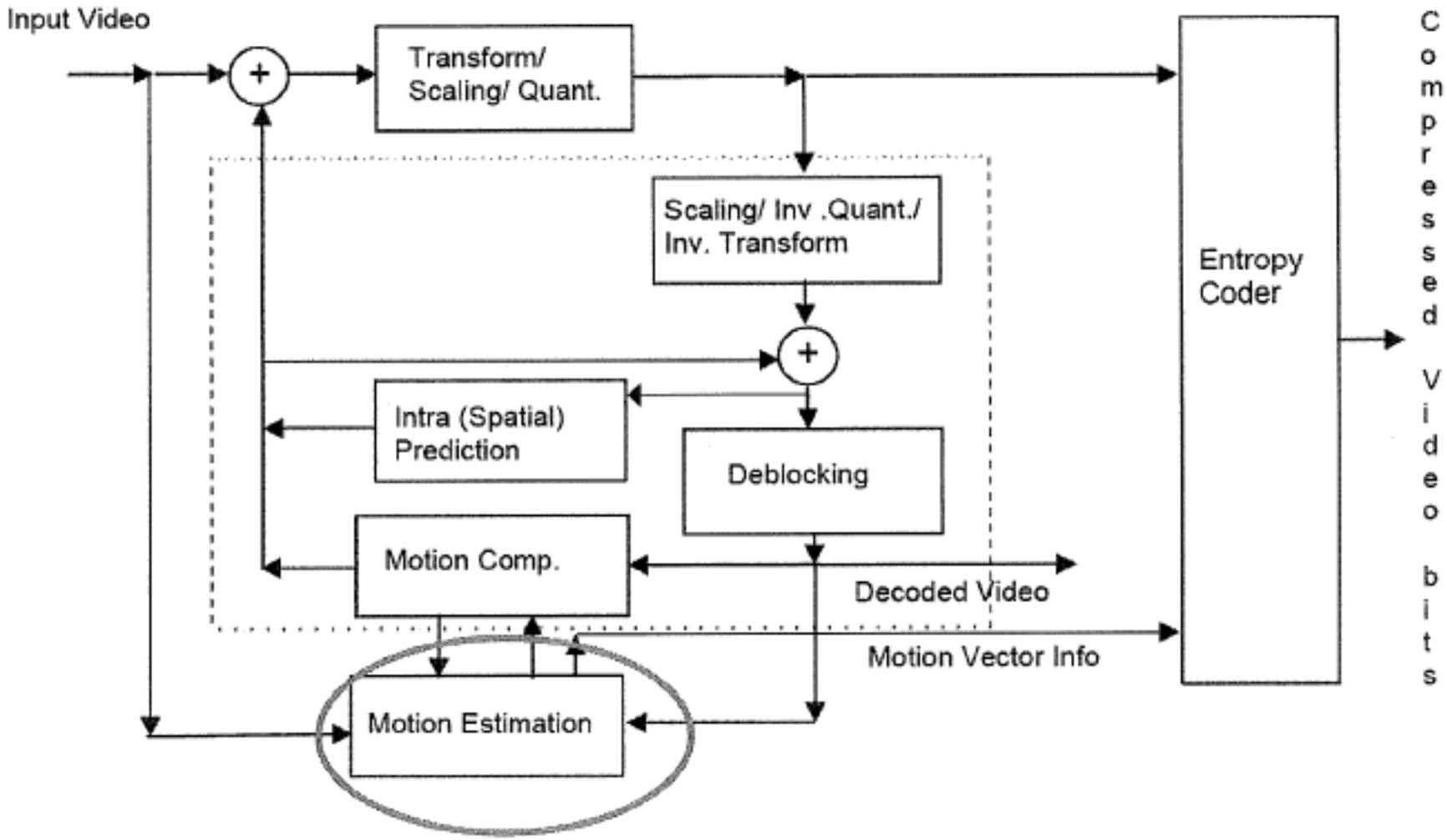
Video encoding standards

- New video encoding standards: H264
- Also known as MPEG-4/ Part 10/Avc
- And H.264 / AVC (Advanced video encoding)
- Provides Higher encoding efficiency
  - 50% better than MPEG-2
  - 30% better than MPEG-4/H264
- Default standard for
  - Blue-ray disks
  - HDTV broadcasts
  - Streaming video in broadcasting
  - Apps in mobile and portable devices
  - Supports resolutions up to 8K UHD (Ultra High Density)

Main Features

- Integer transform in 4x4 blocks. Low complexity, no drifting.
- Variable block size motion compensation 16x16 -> 4x4
- Quarter pixel accuracy in motion vectors (via interpolation)
- Multiple reference picture motion compensation
- Directional spatial prediction for intra frames
- In loop deblocking filtering
- Context adaptive variable length coding (CAVLC)
- Context adaptive binary arithmetic coding (CABAC)
- Robustness to data errors, data losses, synchronization

Motion Estimation



H.264/AVC Encoder [2]

(from: Rahul Vanan, University of Washington)

The H.264 provides asymmetric coding. It is rather a family of standards providing guidelines rather than specific algorithms for implementation. Among these guidelines is the format for encoding and decoding data under different number of profiles. A large number of algorithms ranging in complexity and sophistication have emerged by numerous vendors.
A detailed description of such algorithms are beyond the objectives of this course.

http://helios.mi.parisdescartes.fr/~lomn/Cours/CV/SeqVideo/H.265-HEVC-Tutorial-2014-ISCAS.pdf

Design and Implementation of Next Generation Video Coding Systems (H.265/HEVC Tutorial)

Vivienne Sze (sze@mit.edu)
Madhukar Budagavi (m.budagavi@samsung.com)

ISCAS Tutorial 2014

Instructors

- Vivienne Sze (Assistant Professor at MIT)
  - Involved with video implementation research and standards for 7+ years
    - Contributed over 70 technical documents to HEVC.
    - Within JCT-VC Committee, Primary Coordinator of the core experiments on coefficient scanning and coding; chairman of ad hoc groups on topics related to entropy coding and parallel processing.
    - Published over 25 journal and conference papers.
- Madhukar Budagavi (Research Director at Samsung Research America)
  - Involved with video standards and product development for 15+ years
    - Contributed over 100 technical documents to HEVC.
    - Within JCT-VC Committee, Chaired and co-chaired sub-group activities on spatial transforms, quantization, entropy coding, in-loop filtering, intra prediction, screen content coding and scalable HEVC (SHVC).
    - Published over 40 journal and conference papers, book chapters.

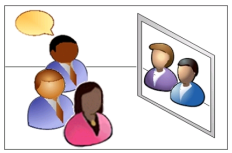
Outline of Tutorial

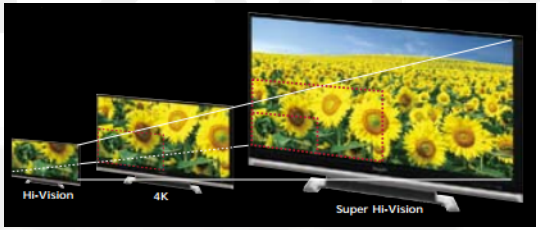
- Part I: Overview of current video coding technology and systems
- Part II: High Efficiency Video Coding (HEVC)
- Part III: Video Codec Implementations
- Part IV: Emerging Applications and HEVC Extensions

Part I: Overview of current video coding technology and systems

Growing Demand for Video

- Video exceeds half of internet traffic and will grow to 86 percent by 2016. Increase in applications, content, fidelity, etc. → Need higher coding efficiency!
- Ultra-HD 4K broadcast expected for Japan in 2014. London Olympics Opening and Closing Ceremonies shot in Ultra-HD 8K. → Need higher throughput!
- 25x increase in mobile data traffic over next five years. Video is a "must have" on portable devices. → Need lower power!

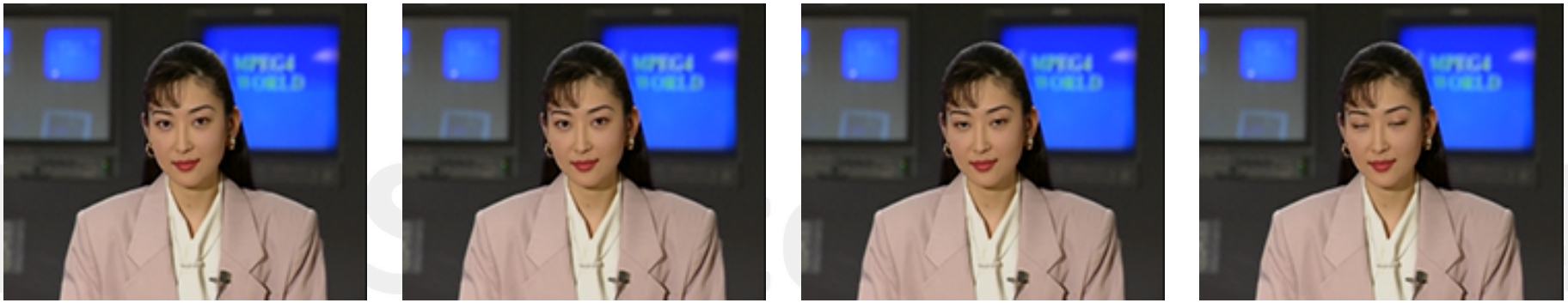


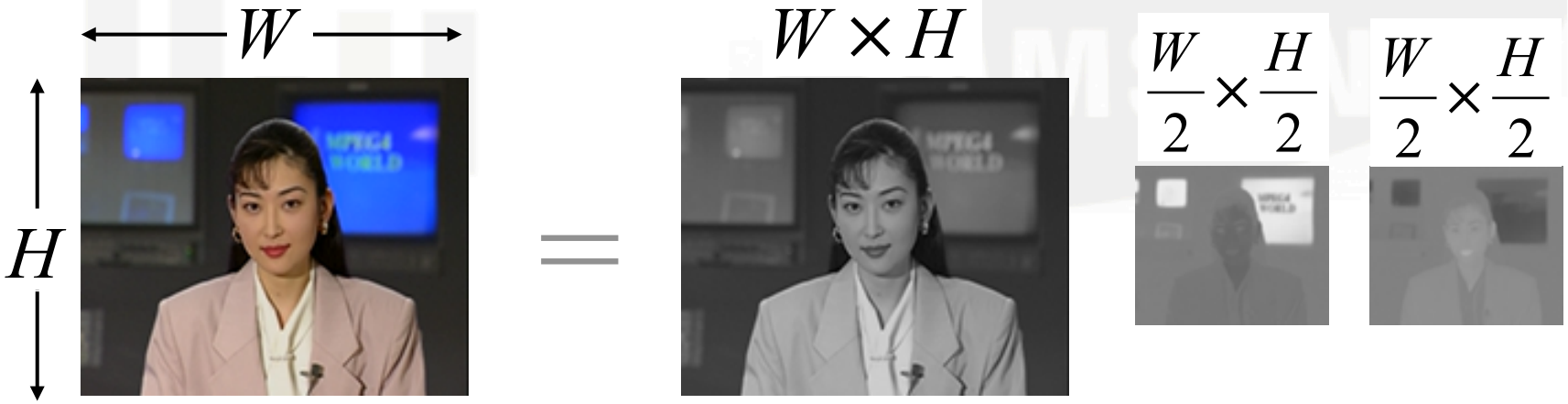






Sources: Cisco Visual Networking Index
Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update

Digital Video





Video Compression

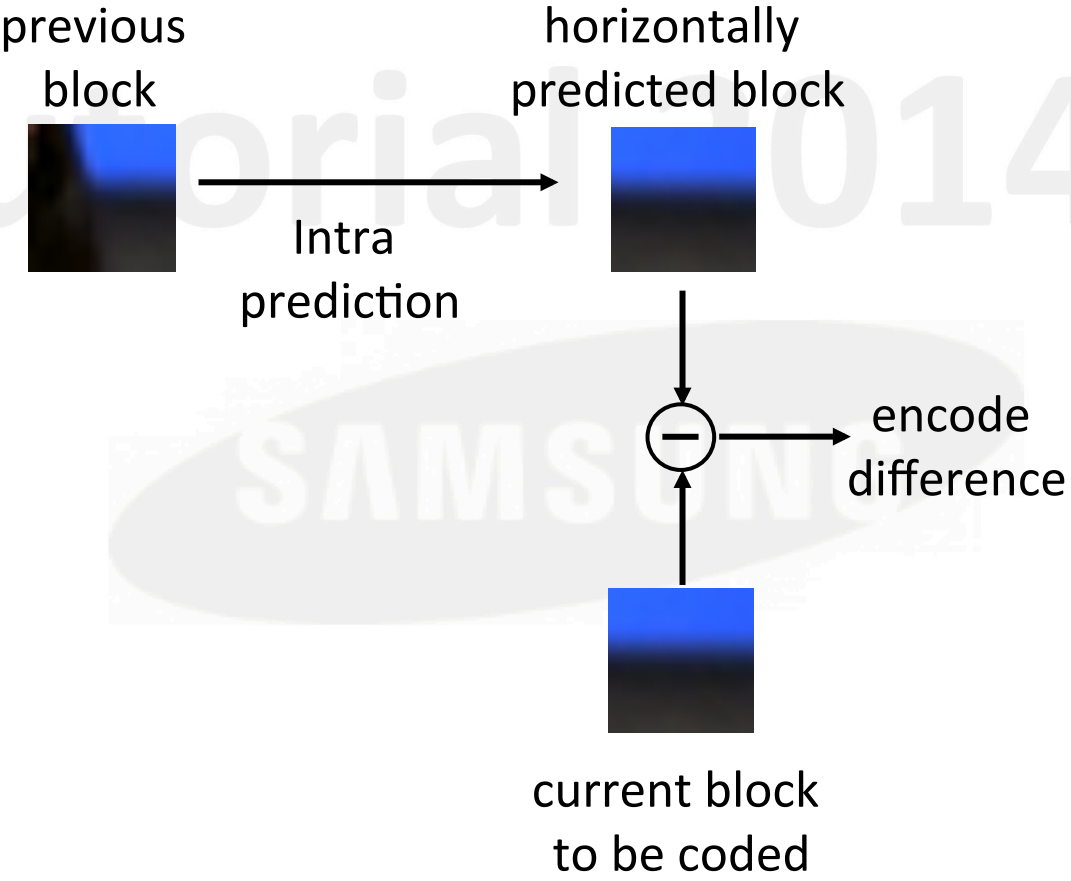
- Uncompressed 1080p high definition (HD) video at 24 frames/ second
  - Pixels per frame: 1920x1080
  - Bits per pixel: 8-bits x 3 (RGB)
  - 1.5 hours: 806 GB
  - Bit-rate: 1.2 Gbits/s
- Blu-Ray DVD
  - Capacity: 25 GB (single layer)
  - Read rate: 36 Mbits/s
- Video Streaming or TV Broadcast
  - 1 Mbits/s to 20 Mbits/s
- Require 30x to 1200x compression

Video Compression Basics

- Compression is achieved by removing redundant information from the video sequence
- Types of redundancies in video sequences
  - Spatial redundancy
  - Perceptual redundancy
  - Statistical redundancy
  - Temporal redundancy

Spatial Redundancy Removal (1)

- Intra prediction



Spatial Redundancy Removal (2)

- Block Transforms
  - Typically matrix operations
  - Used for correlation reduction and energy compaction in the block

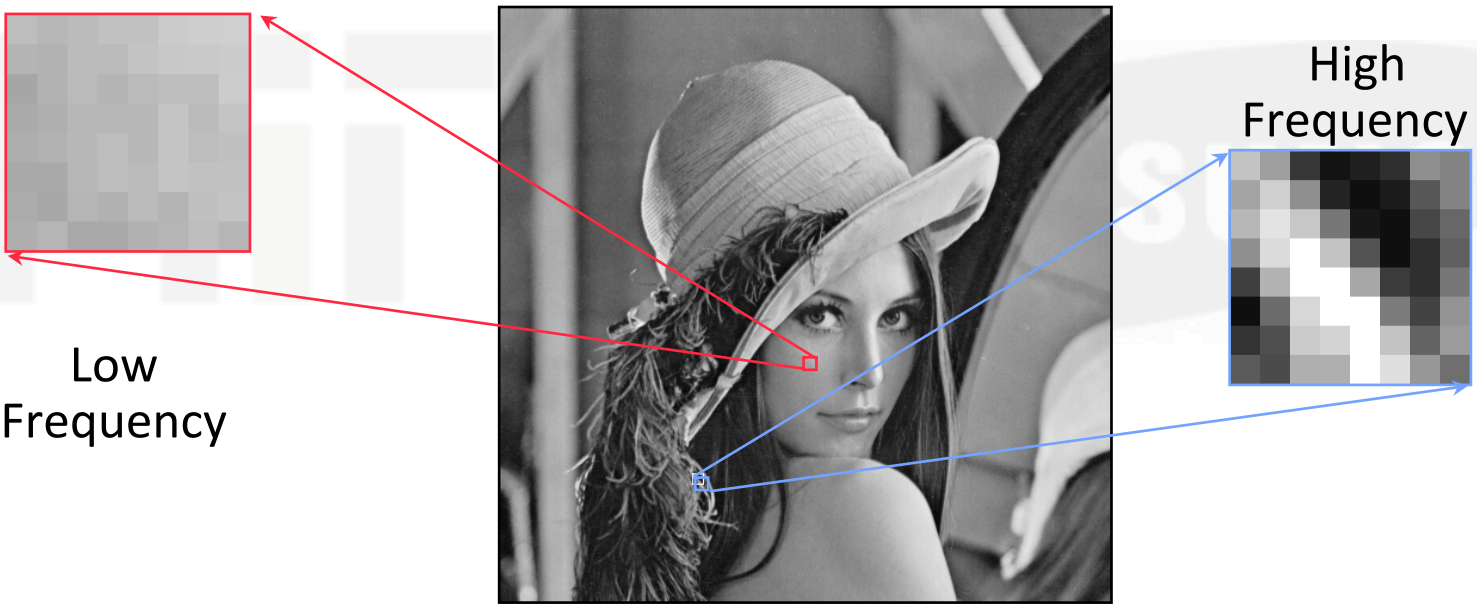
151	149	145	140	136	133	128	120
150	147	144	140	136	132	127	118
149	145	142	138	135	129	122	116
147	143	139	136	131	126	120	113
141	139	137	132	127	124	116	109
138	135	133	130	125	120	113	106
135	131	130	128	123	117	111	105
132	130	129	126	120	115	109	105

8x8 2D Discrete Cosine Transform (DCT)

1037	80	0	9	0	4	0	0
49	1	3	3	0	0	0	1
0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	0
0	0	1	0	0	0	0	0
1	1	1	1	2	0	0	0
0	1	0	0	0	0	0	0
0	0	0	0	0	0	1	0

Perceptual Redundancy Removal (1)

- Not all video data are equally significant from a perceptual point of view
- Make use of the properties of the Human Visual System (HVS)
  - HVS is more sensitive to low frequency information



Perceptual Redundancy Removal (2)

- Quantization is a good tool for perceptual redundancy removal
  - Most significant bits (MSBs) are perceptually more important than least significant bits (LSBs)
  - Coefficient dropping (quantization with zero bits) example:

Original frame

Image obtained by retaining 36 DCT coefficients for each 8x8 block

Statistical Redundancy Removal (1)

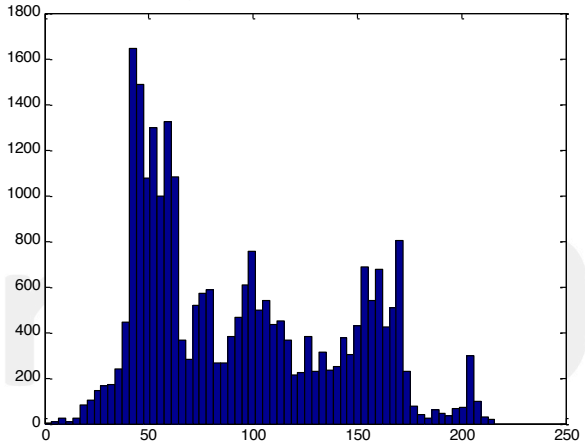
- Not all pixel values in an image (or in the transformed image) occur with equal probability

- Use entropy coding (e.g. variable length coding)
  - Shorter codewords used to represent more frequent values
  - Longer codewords used to represent less frequent value

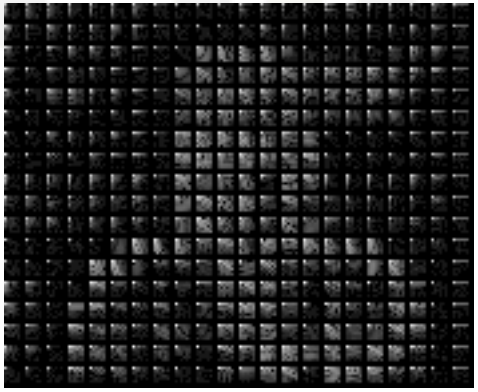
Statistical Redundancy Removal (2)

- Original image: 8 bits/pixel, Entropy coding: 7.14 bits/pixel

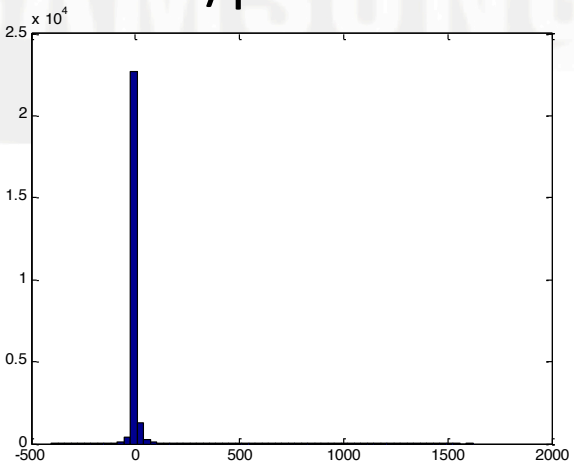
Histogram



- Results more dramatic when entropy coding is applied on transformed and quantized image: 1.82 bits/pixel

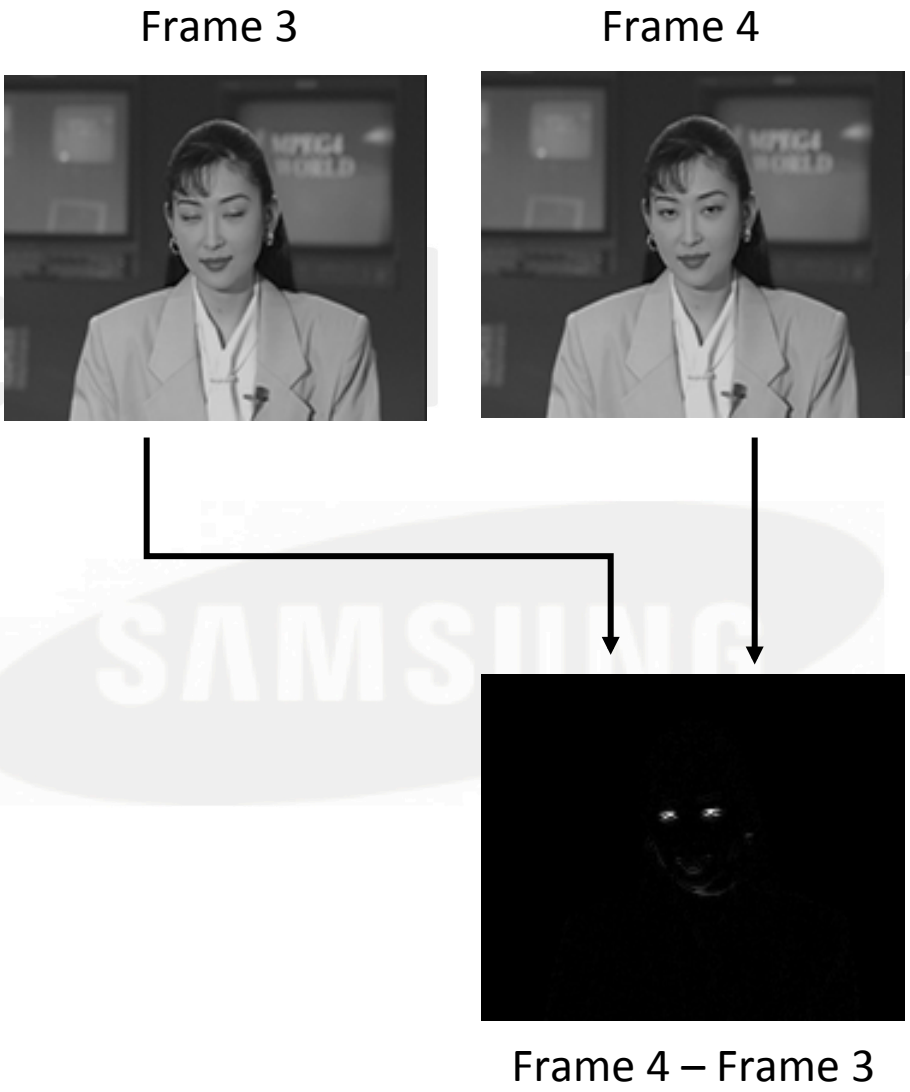


Histogram



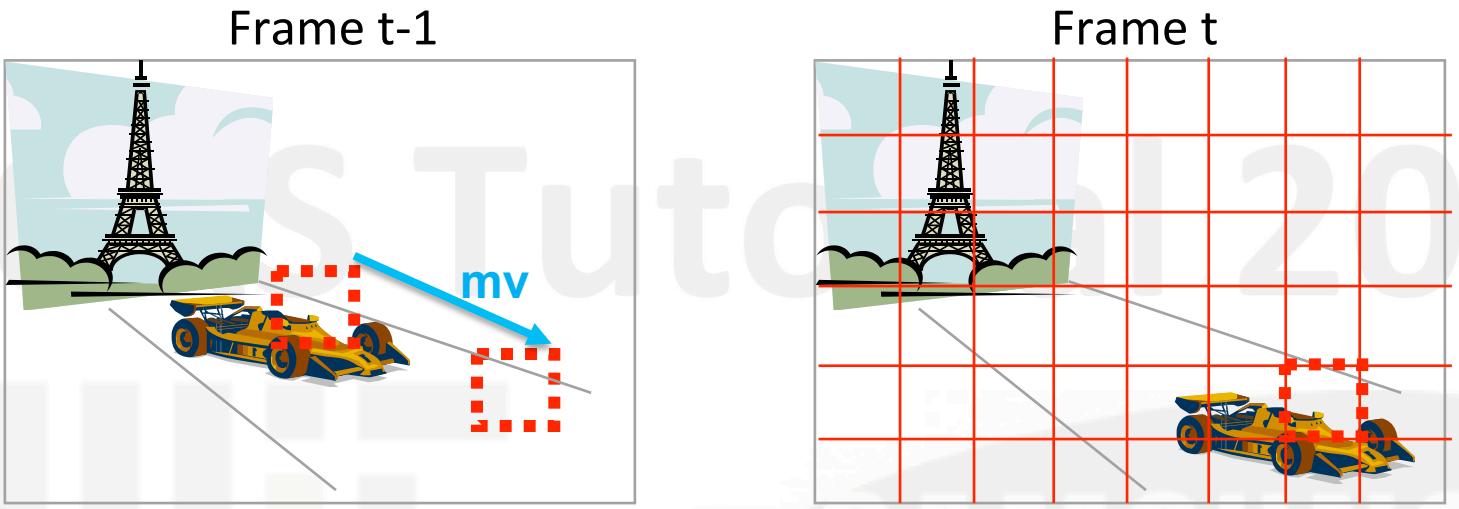
Temporal Redundancy Removal (1)

- Inter prediction
- Frame difference coding
  - Difference can be encoded using DCT + Quantization + Entropy Coding



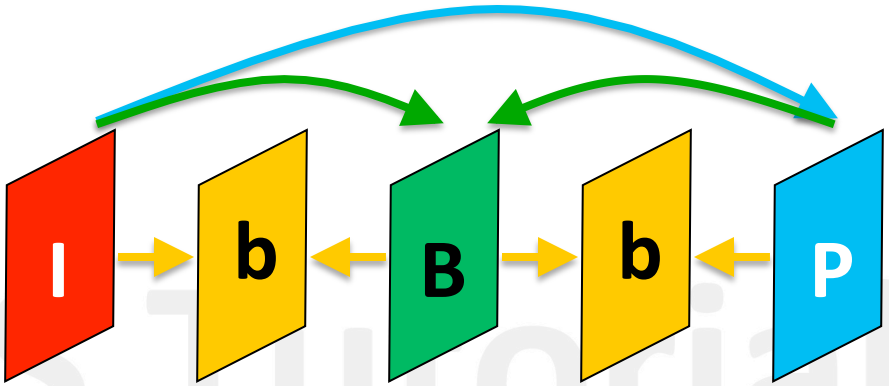
Temporal Redundancy Removal (2)

- Inter prediction using Motion compensated prediction



  - Divide the frame into blocks and apply block motion estimation/ compensation
  - For each block find out the relative motion between the current block and a matching block of the same size in the previous frame
  - Transmit the motion vector(s) for each block

Temporal Prediction and Picture Coding Types

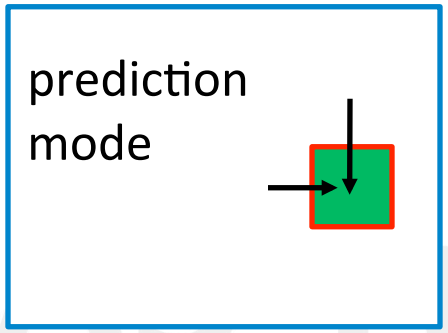


- Intra Picture (I)
  - Picture is coded without reference to other pictures
- Inter picture (P, B, b)
  - Uni-directionally predicted (P) Picture
    - Picture is predicted from one prior coded picture
  - Bi-directionally predicted (B, b) Picture
    - Picture is coded from two prior coded pictures

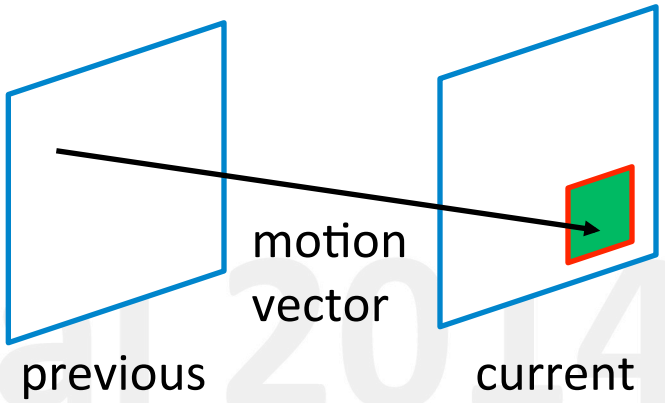
Summary of Key Steps in Video Coding

- Intra Prediction and Inter Prediction

Intra Predication

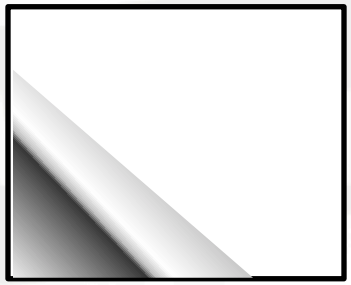


Inter Prediction (Motion Compensation)



- Transform and Quantization of residual (prediction error)

many pixels*



few coefficients

- Entropy coding on syntax elements

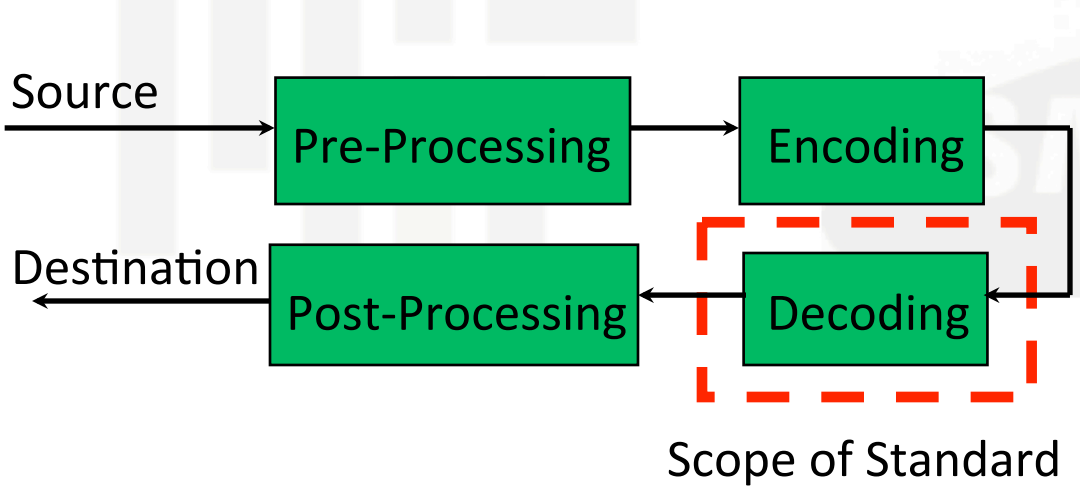
  e.g. prediction modes, motion vectors, coefficients

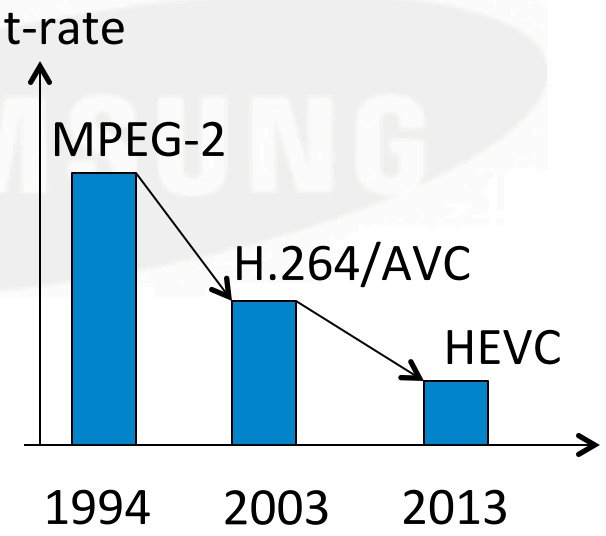
- In-loop filtering to reduce coding artifacts

* Residual figure from J. Apostolopoulos, "Video Compression," MIT 6.344 Lecture, Spring 2004

Video Compression Standards

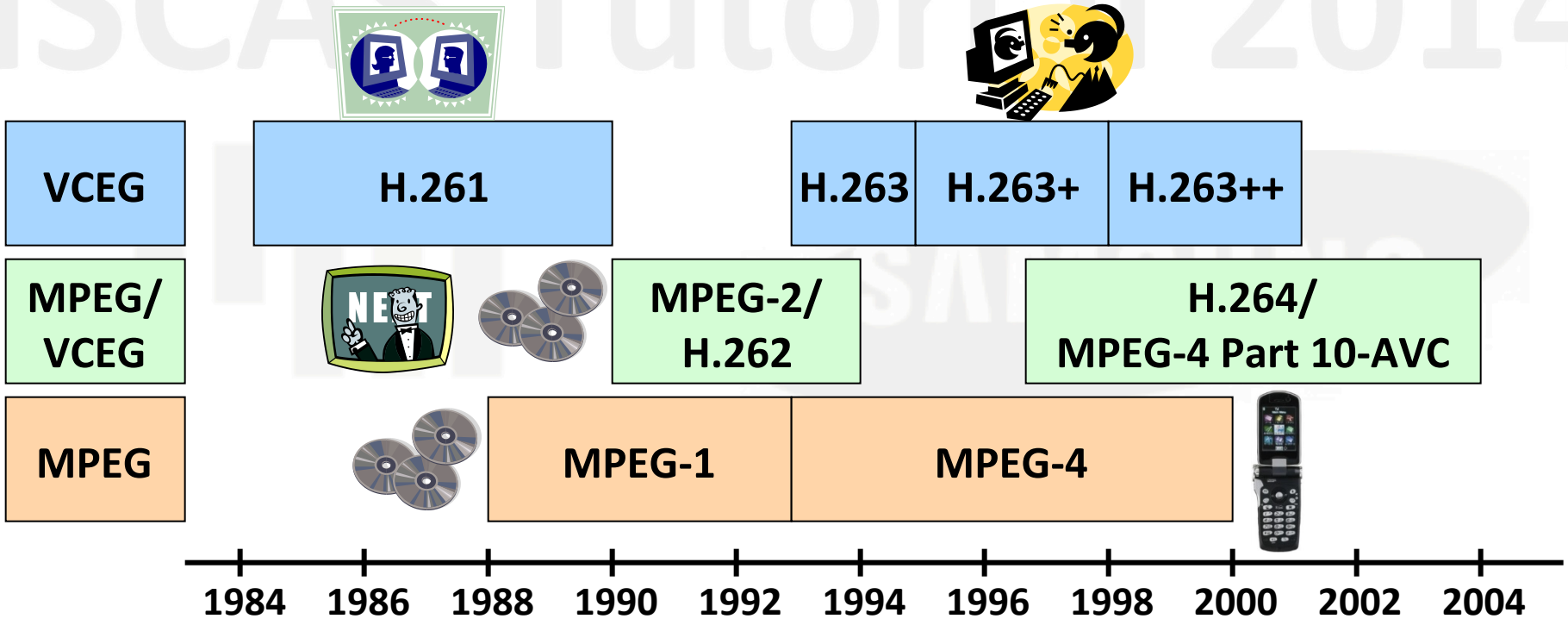
- Ensures inter-operability between encoder and decoder
- Support multiple use cases and applications
  - Levels and Profiles
- Video coding standard specifies decoder: mapping of bits to pixels
- ~2x improvement in compression every decade



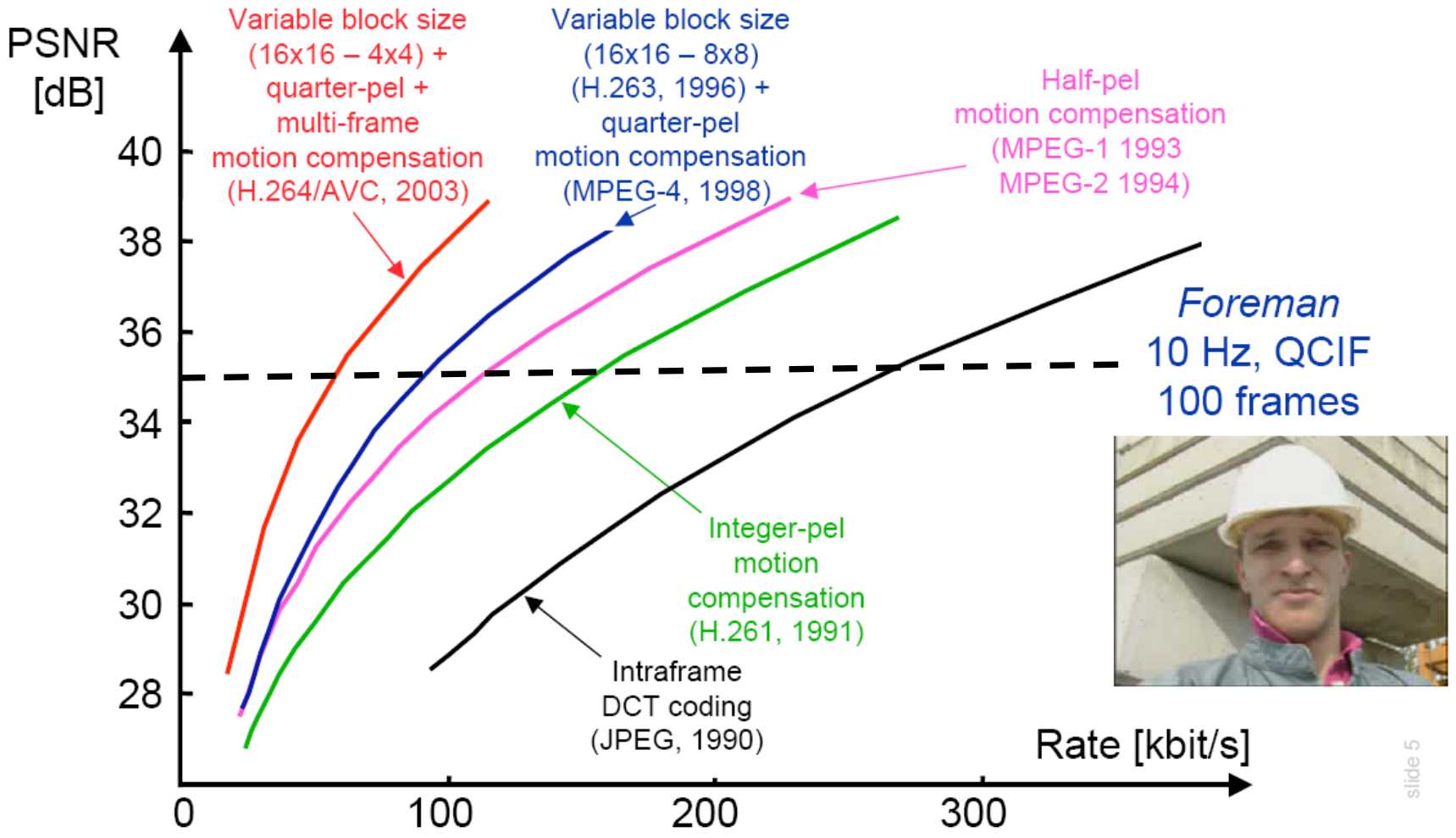


History of Video Coding Standards

- MPEG: Moving Picture Experts Group (ISO/IEC)
- VCEG: Video Coding Experts Group (ITU-T)
- Other standards: VC1, VP8/VP9, China AVS, RealVideo



Video Coding Progress



Source: T. Wiegand, JVT-W132, 2007

H.264/MPEG-4 AVC

- Completed (version 1) in May 2003
- H.264/AVC is the most popular video standard in market
  - 80% of video on the internet is encoded with H.264/AVC
- Applications include
  - HDTV broadcast satellite, cable, and terrestrial
  - video content acquisition and editing
  - camcorders, security applications, Internet and mobile network video, Blu-ray Discs
  - real-time video chat, video conferencing, and telepresence
- ~50% higher coding efficiency than MPEG-2 (used in DVD, US terrestrial broadcast)

Improvements of H.264/MPEG-4 AVC over previous standards

- Prediction
  - Intra prediction using neighboring samples
  - Temporal prediction using multiple frames
  - Motion compensation on variable block size, quarter-pel
- Transform
  - 4x4/8x8 Integer transform, 2x2/4x4 Secondary Hadamard
- Quantization
  - Finer quantization supported
- Entropy coding
  - Context adaptive variable length coding (CAVLC) and arithmetic coding (CABAC)
- In-loop deblocking filter

Part II: High Efficiency Video Coding (HEVC)

High Efficiency Video Coding (HEVC)

- Achieves 2x higher compression compared to H.264/AVC
- High throughput (Ultra-HD 8K @ 120fps) & low power
  - Implementation friendly features (e.g. built-in parallelism)
- Benefits include
  - reduce the burden on global networks
  - easier streaming of HD video to mobile devices
  - account for advancing screen resolutions (e.g. Ultra-HD)

"HEVC will provide a flexible, reliable and robust solution, future-proofed to support the ***next decade of video****"*

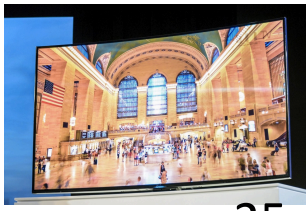
- ITU-T Press Release (2013)

Samsung Galaxy S4

Netflix Ultra-HD 4K

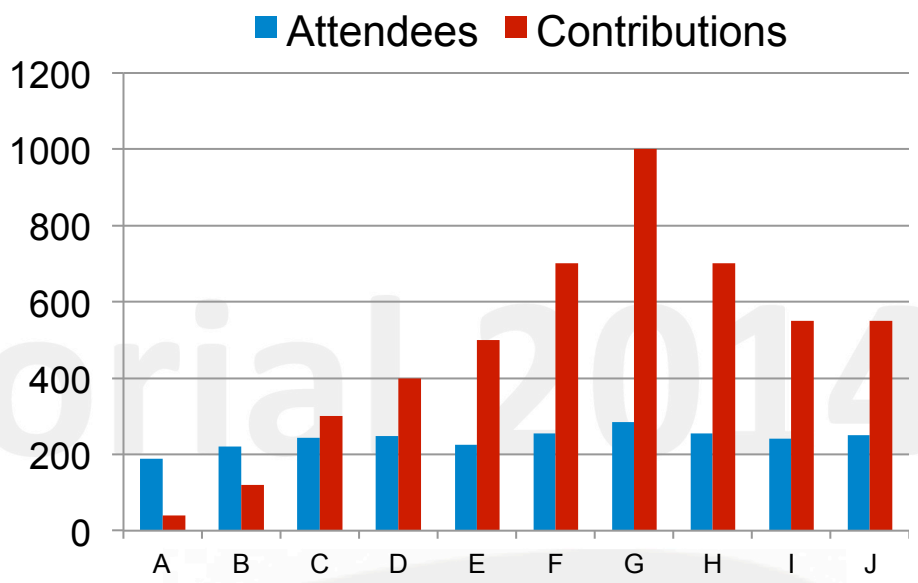
Live delivery of French Open

Samsung TV Ultra-HD 4K



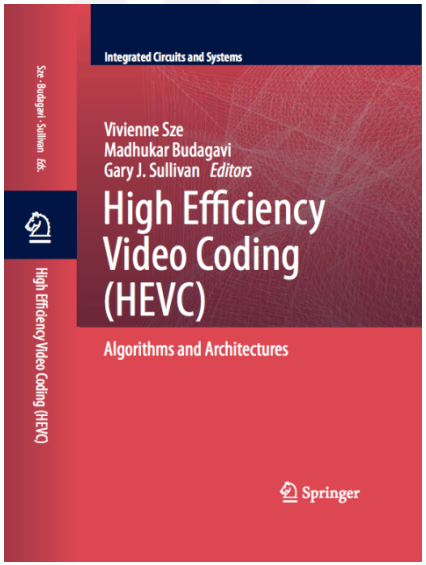
Activity in JCT-VC Committee

- Chairs
  - G. J. Sullivan (Microsoft)
  - J. R. Ohm (Aachen University)
- Meet Quarterly
  - 1st meeting (A) [January 2010]
  - .....
  - 12th meeting (L) [January 2013]
- ~250 attendees per meeting representing ~70 companies
- Several hundred contributions per meeting
- Each meeting is around 9 - 10 days (14+ hours/day)
- Multiple parallel tracks



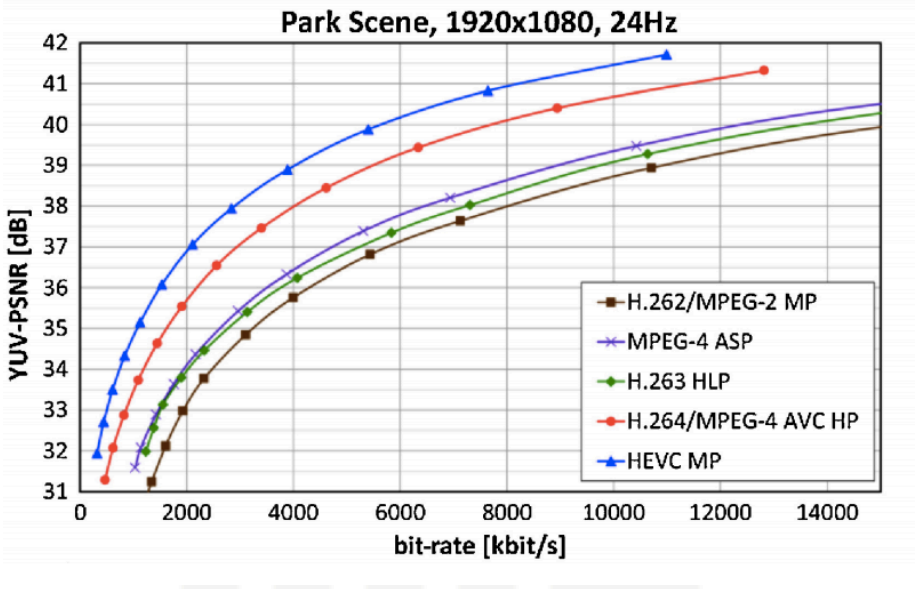
HEVC Reference Documents

- Meeting Contributions
  - http://phenix.int-evry.fr/jct/
- Specification
  - http://www.itu.int/ITU-T/recommendations/rec.aspx?rec=11885
- Reference Software (HM)
  - https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/



- References
  - G. J. Sullivan, et al. "Overview of the High Efficiency Video Coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology,* 2012
  - V. Sze, M. Budagavi, G. J. Sullivan (Editors), "High Efficiency Video Coding (HEVC): Algorithms and Architectures," Springer, 2014 http://www.springer.com/engineering/signals/book/978-3-319-06894-7

Coding Efficiency of HEVC (Objective)



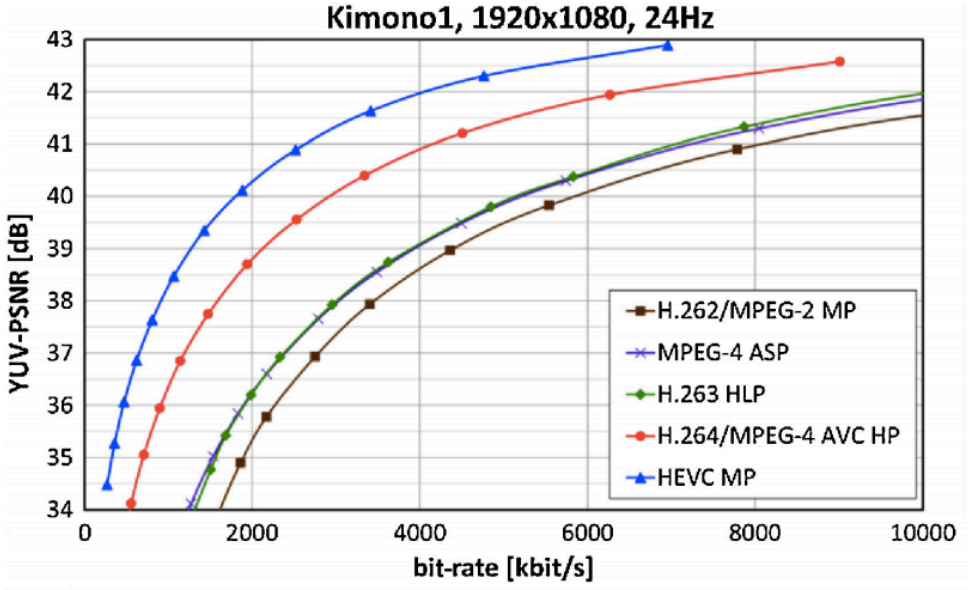


TABLE VI

AVERAGE BIT-RATE SAVINGS FOR EQUAL PSNR FOR ENTERTAINMENT APPLICATIONS

Encoding	Bit-Rate Savings Relative to			
	H.264/MPEG-4 AVC HP	MPEG-4 ASP	H.263 HLP	MPEG-2/ H.262 MP
HEVC MP	35.4%	63.7%	65.1%	70.8%
H.264/MPEG-4 AVC HP	–	44.5%	46.6%	55.4%
MPEG-4 ASP	–	–	3.9%	19.7%
H.263 HLP	–	–	–	16.2%

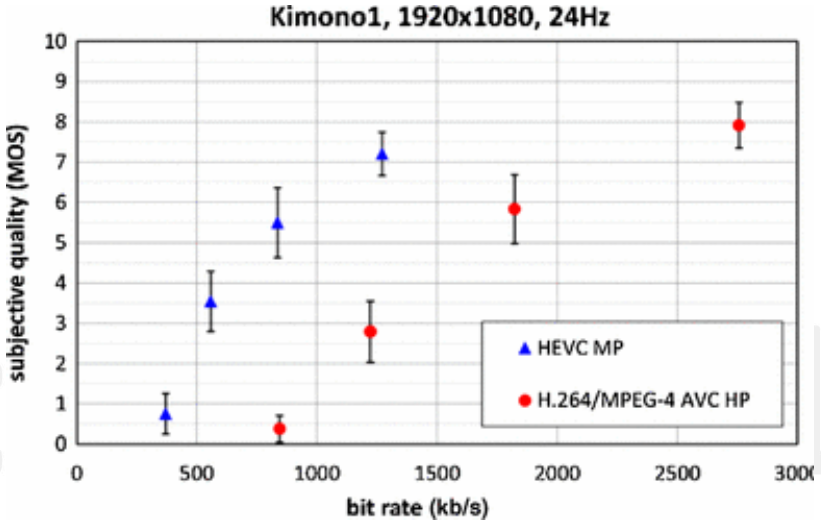
$$PSNR = 10\log_{10}\frac{(2^{bitdepth}-1)^2 * W * H}{\sum_i \{O_i - D_i\}^2}$$

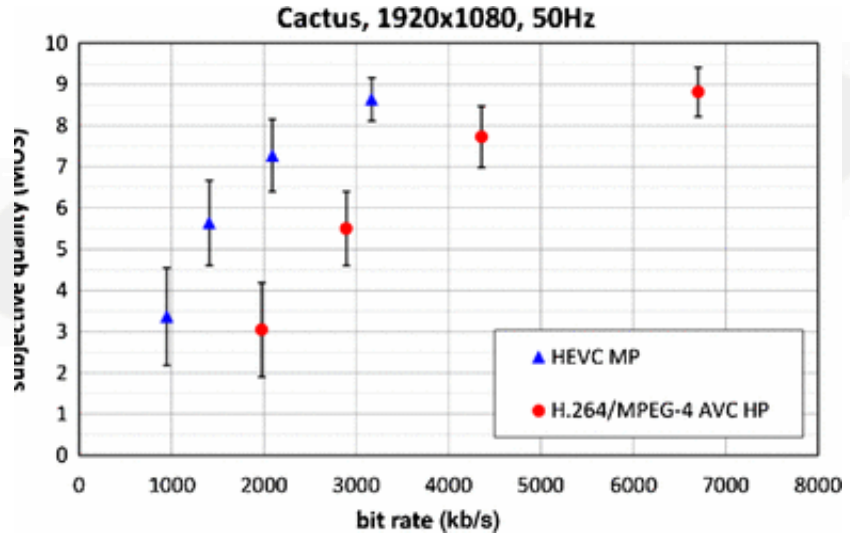
J. R. Ohm et al., "Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC),"*IEEE Transactions on Circuits and Systems for Video Technology*, 2012

Coding Efficiency of HEVC (Subjective)

Subjective Tests for Entertainment Applications (Random Access)

Sequences	Bit-rate Savings
BQ Terrace	63.1%
Basketball Drive	66.6%
Kimono1	55.2%
Park Scene	49.7%
Cactus	50.2%
BQ Mall	41.6%
Basketball Drill	44.9%
Party Scene	29.8%
Race Horse	42.7%
Average	**49.3%**

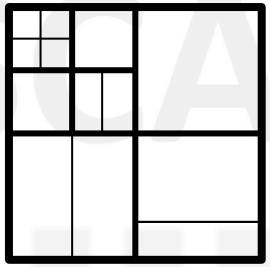




J. Ohm et al., "Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC),"*IEEE Transactions on Circuits and Systems for Video Technology*, 2012

H.265/HEVC vs. H.264/AVC Decoder

64x64



Larger and Flexible Coding Block Size

Larger Interpolation Filter

Fewer Edges

High Throughput CABAC & Advanced Motion Vector Prediction

Larger Transforms and More Sizes

More Prediction Modes

Encoded bitstream

Entropy Decoder

Motion Comp.

Intra Prediction

Q^{-1} $+T^{-1}$

Picture Buffer

Decoded pixels

Sample Adaptive Offset

Deblocking Filter

In-loop Filter

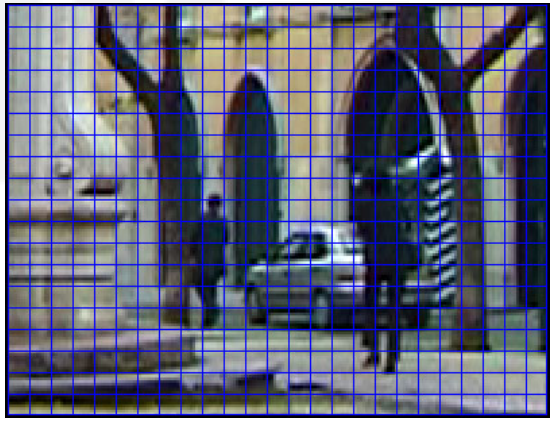
Key Features In HEVC

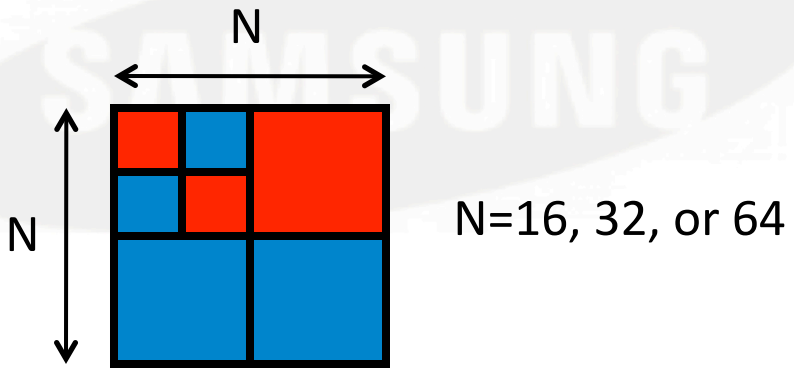
	High Coding Efficiency	High Throughput / Low Power
Larger and Flexible Coding Block Size	X	
More Sophisticated Intra Prediction	X	
Larger Interpolation Filter for Motion Compensation	X	
Larger Transform Size	X	
Parallel Deblocking Filter		X
Sample Adaptive Offset	X	
High Throughput CABAC	X	X
High Level Parallel Tools		X
Parallel Merge/Skip		X

M. Zhou, V. Sze, M. Budagavi, "Parallel Tools in HEVC for High-Throughput Processing," *SPIE Optical Engineering + Applications, Applications of Image Processing XXXV*, 2012.

Larger Coding Blocks

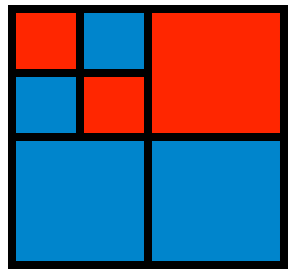
- Each frame is broken up into blocks
- Large block sizes reduce signaling overhead
- In H.264/AVC, macroblock is always 16x16 pixels
  - Each macroblock is either inter or intra coded
- In HEVC, Coding Tree Unit (CTU) can have up to 64x64 pixels
  - CTU can have a combination of inter and intra coded blocks

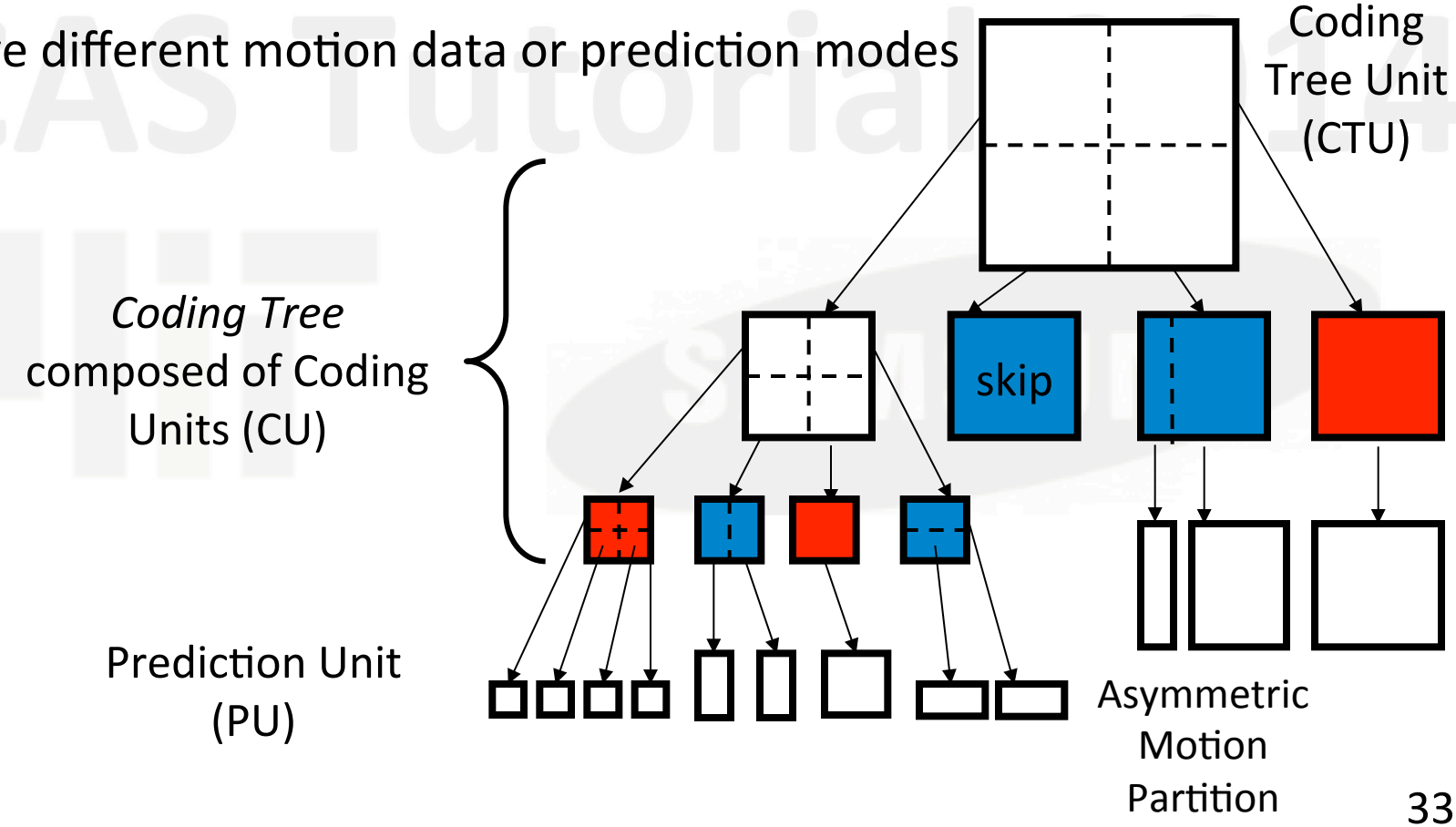




Flexible Coding Block Structure

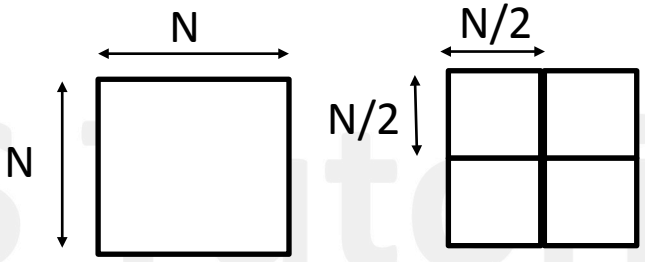
- Better adaptation to different video content
- CTU divided into Coding Units (CU) with Quad tree
- Coding units divided into prediction units (PU)
- PU have different motion data or prediction modes





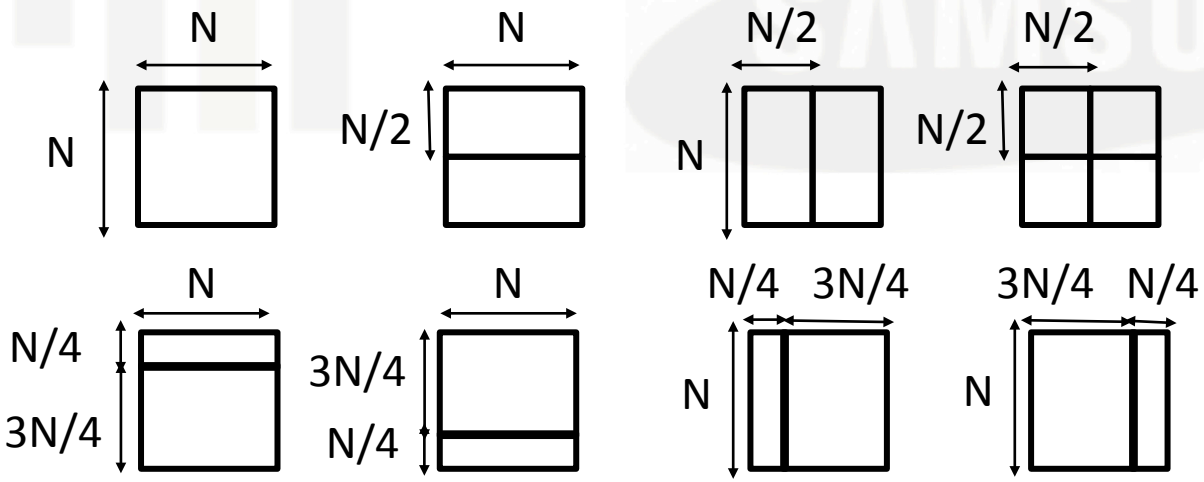
Prediction Units

- Intra-Coded CU can only be divided into square partition units
  - For a CU, make decision to split into four PU (8x8 CUs only) or single PU



Two methods of partitioning for intra-coded CU

- Inter-Coded CU can be divide into square and non-square PU as long as one side is at least 4 pixels wide (note: no 4x4 PU)

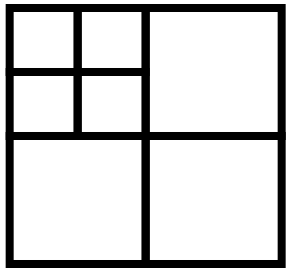


Eight methods of partitioning for inter-coded CU

Large Transforms

- HEVC supports 4x4, 8x8, 16x16, 32x32 integer transforms
  - Two types of 4x4 transforms (IDST-based for Intra, IDCT-based for Inter); IDCT-based transform for 8x8, 16x16, 32x32 block sizes
  - Integer transform avoids encoder-decoder mismatch and drift caused by slightly different floating point representations.
  - Parallel friendly matrix multiplication/partial butterfly implementation
  - Transform size signaled using Residual Quad Tree
- Achieves 5 to 10% increase in coding efficiency
- Increased complexity compared to H.264/AVC
  - 8x more computations per coefficient
  - 16x larger transpose memory

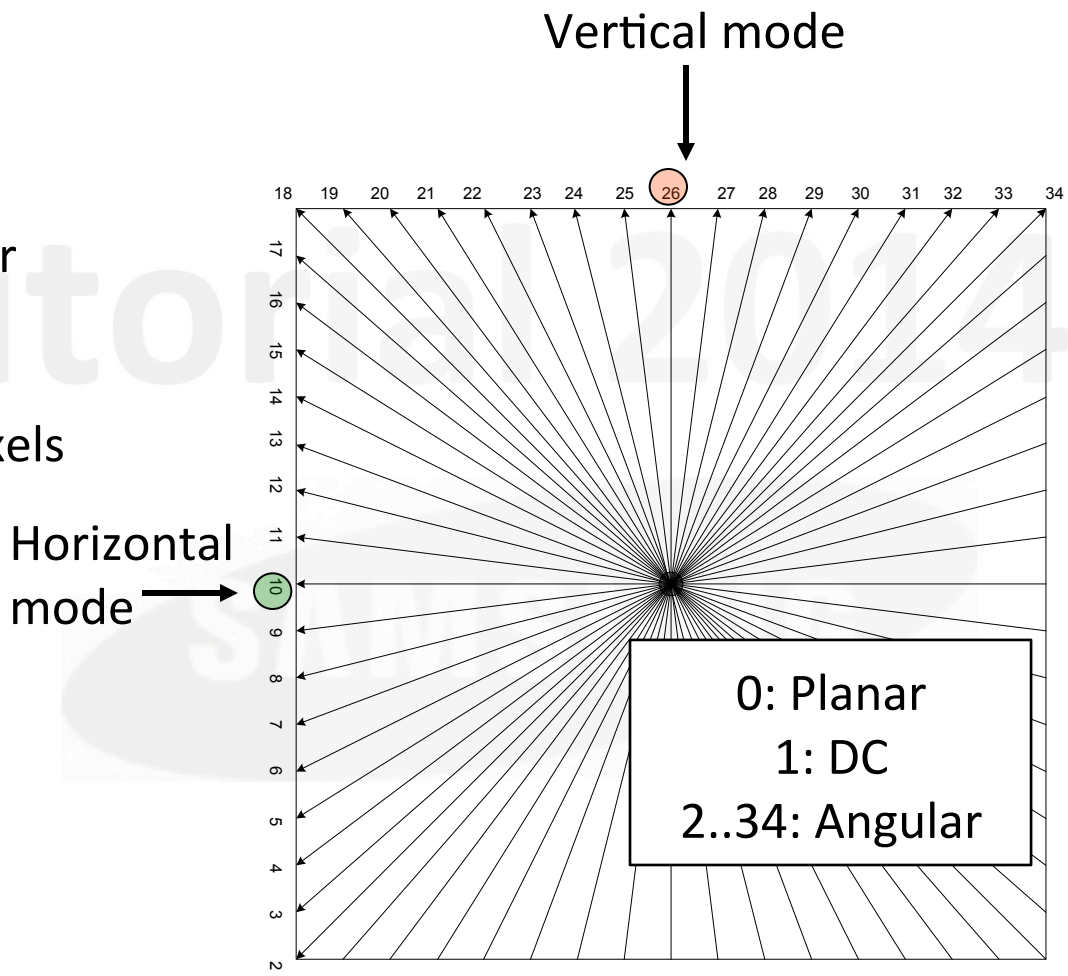
Represent residual of CU with TU quad tree



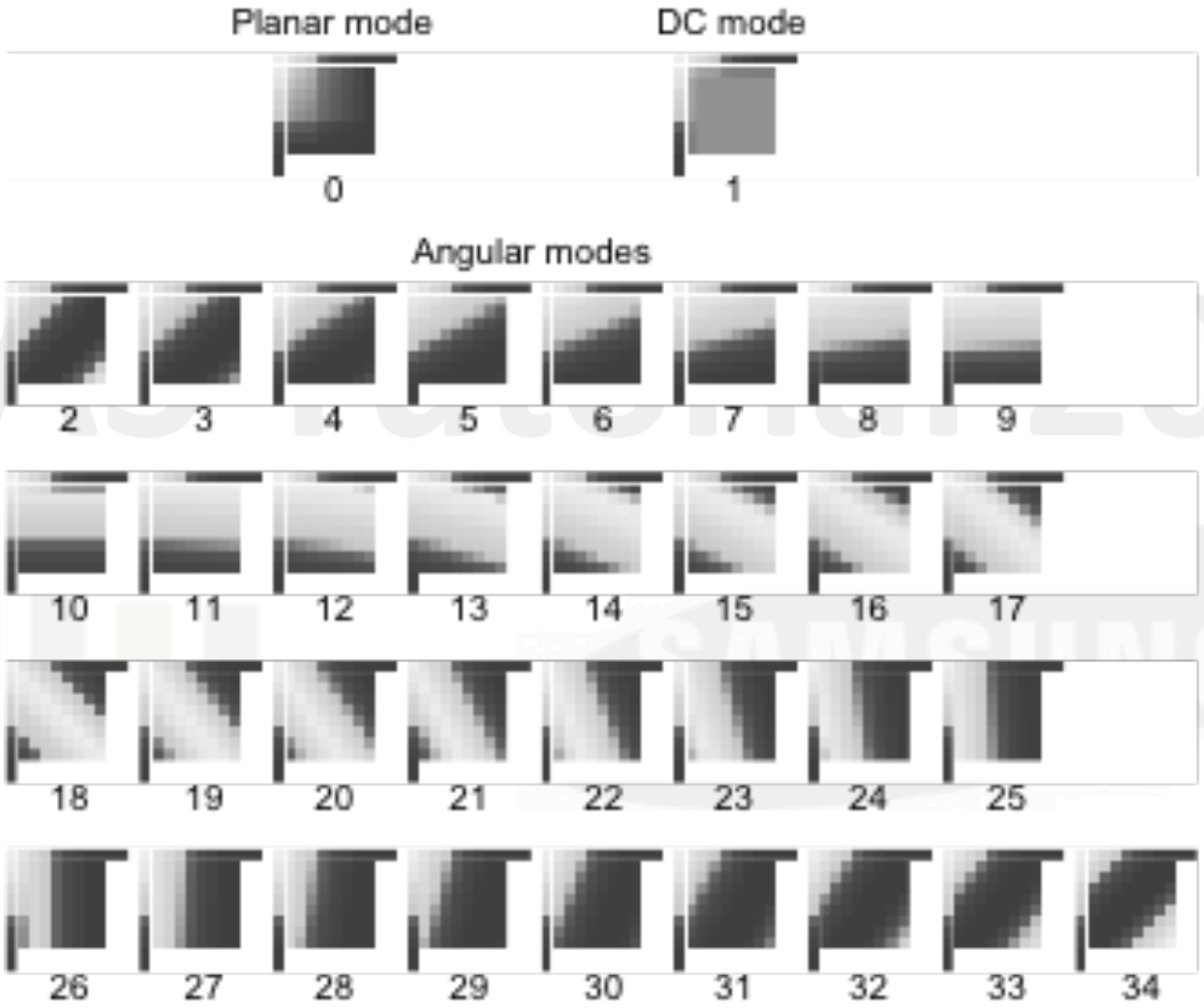
M. Budagavi et al., "Core Transform Design in the High Efficiency Video Coding (HEVC) Standard," IEEE JSTSP, 2013

Intra Prediction

- H.264/AVC has 10 modes
  - angular (8 modes), DC, planar
- HEVC has 35 modes
  - angular (33 modes), DC, planar
- Angular prediction
  - Interpolate from reference pixels at locations based on angle
- DC
  - Constant value which is an average of neighboring pixels (reference samples)
- Planar
  - Average of horizontal and vertical prediction



Intra Prediction Modes



J. Lainema, W.-J. Han, "Intra Prediction in HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014.

Removing Intra Artifacts (Pre-Processing)

- Reference Sample Smoothing
  - Smooth out neighboring pixels (i.e., reference samples) before using them for prediction
  - Reduce contouring artifacts caused by edges in the reference sample arrays
  - Two modes
    - Three-tap smoothing filter
    - Strong intra smoothing with corner reference pixels
  - Application of smoothing depends on PU size and prediction mode

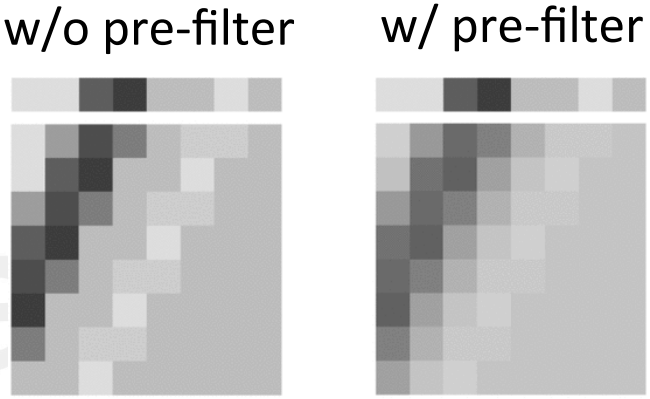
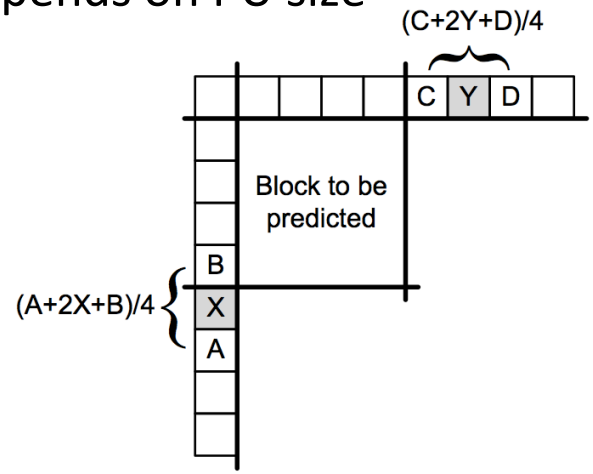
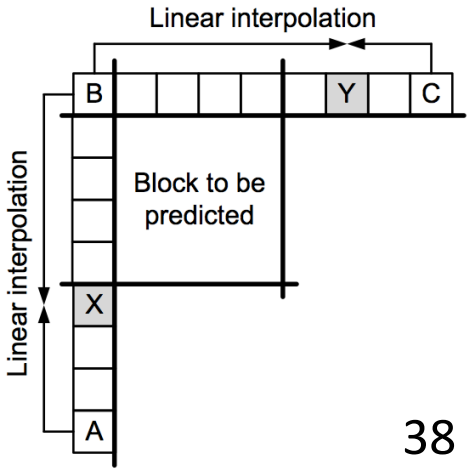


Image source: M. Wien, TCSVT, July 2003

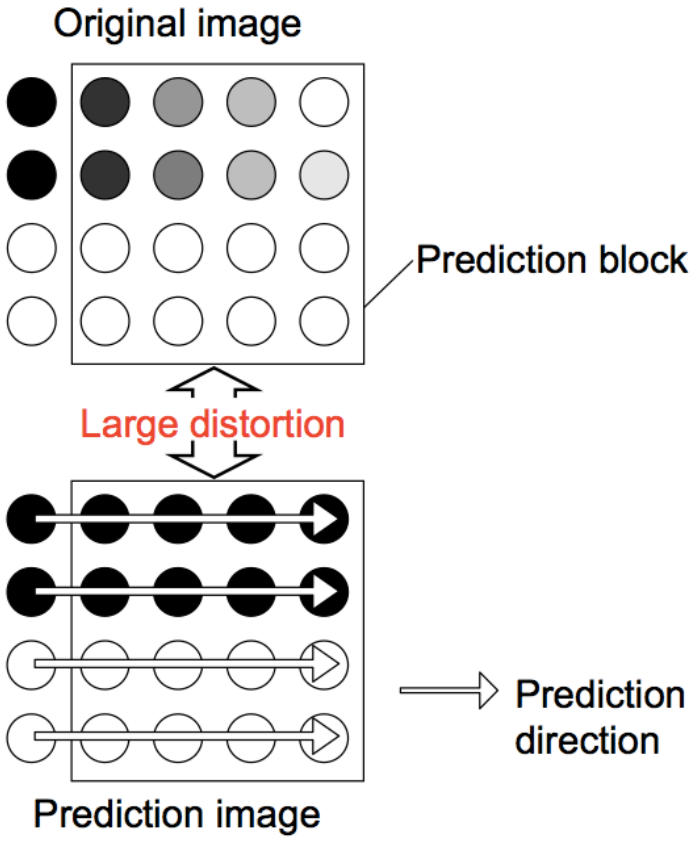




J. Lainema, W.-J. Han, "Intra Prediction in HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014.

Removing Intra Artifacts (Post-Processing)

- Boundary Smoothing
  - Intra prediction may introduce discontinuities along block boundaries
  - Filter first prediction row and column with three-tap filter for DC prediction, and two-tap for horizontal and vertical prediction



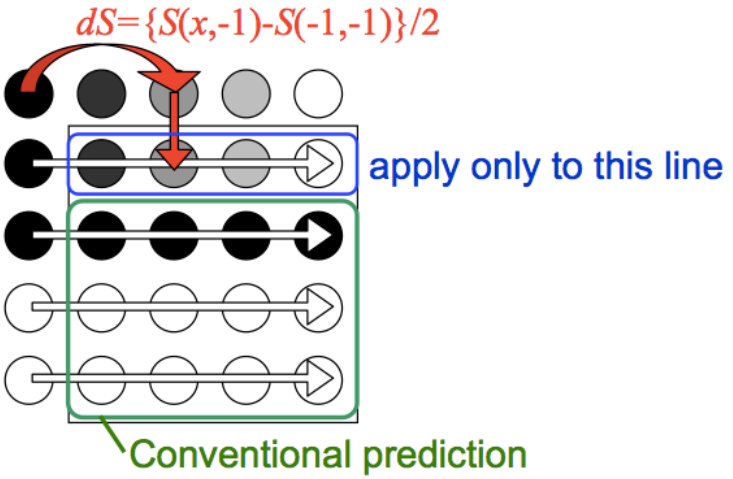
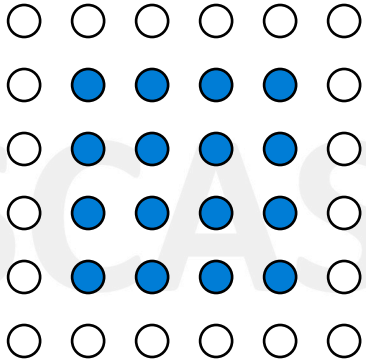


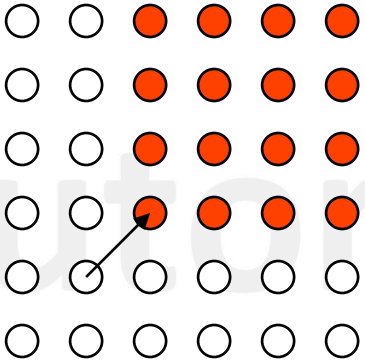
Image source: JCTVC-F172, July 2011

Inter Prediction

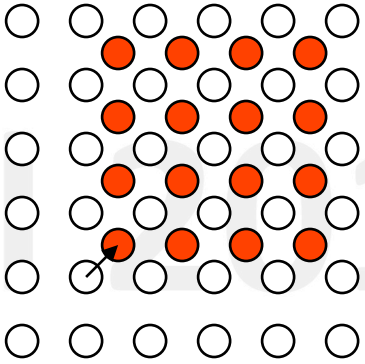
- Motion vectors can have up to ¼ pixel accuracy (interpolation required)



4x4 block in current frame



Reference block in previous frame
Vector (1, -1)



Reference block in previous frame
Vector (0.5, -0.5)

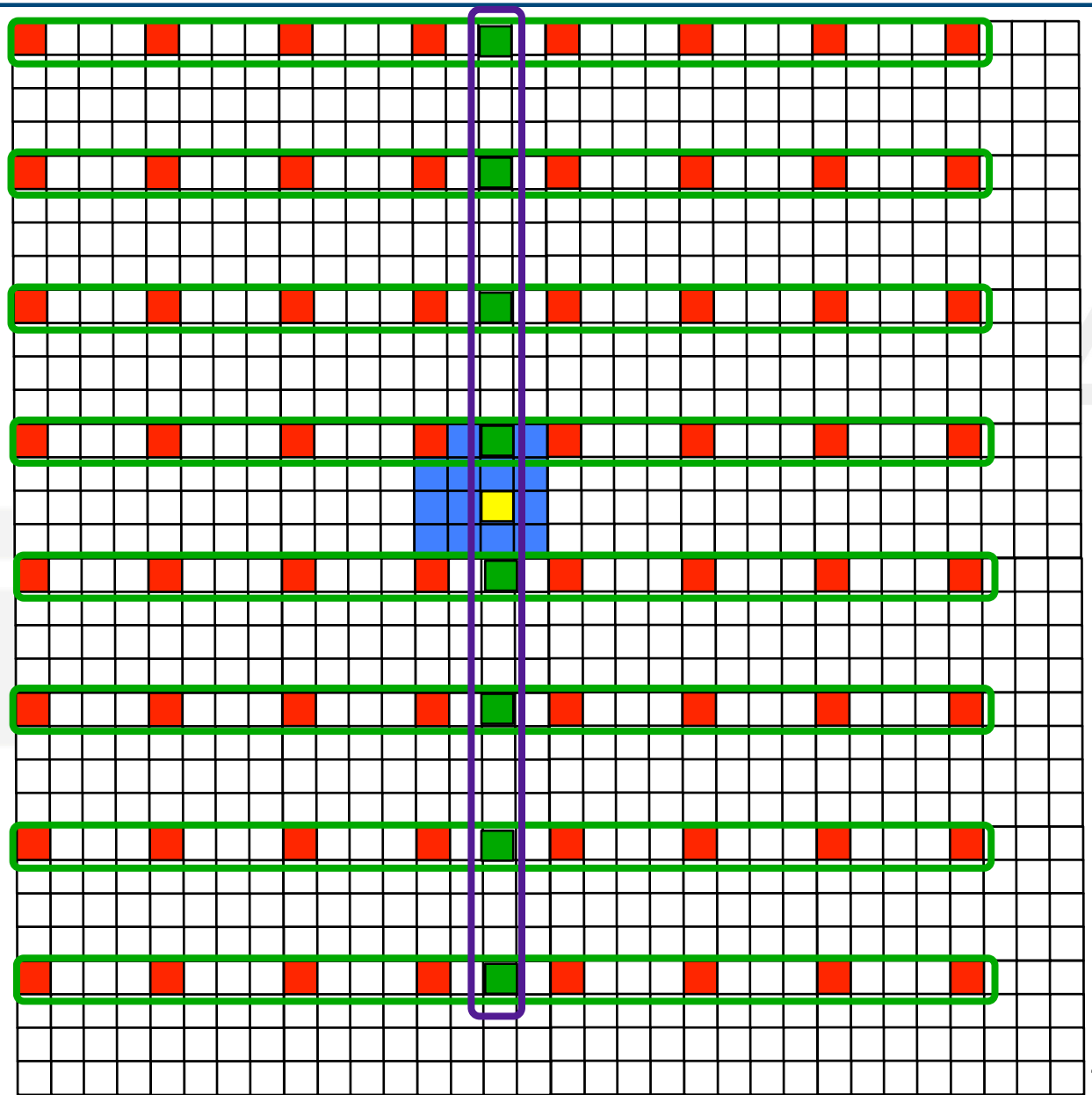
- In H.264/AVC, luma uses 6-tap filter, and chroma uses bilinear filter
- In HEVC, luma uses 8/7-tap and chroma uses 4-tap
  - Different coefficients for ¼ and ½ positions
- Restricted prediction on small PU sizes

Interpolation Filter

Require integer pixels (highlighted in red) to interpolate fractional pixels (highlighted in blue)

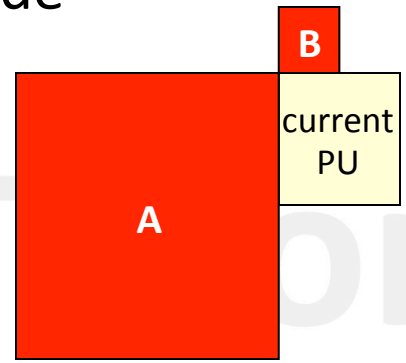
To interpolate NxN pixels requires up to (N+7)x(N+7) reference pixels

Use 1-D filters (order matters for greater than 8-bit video)



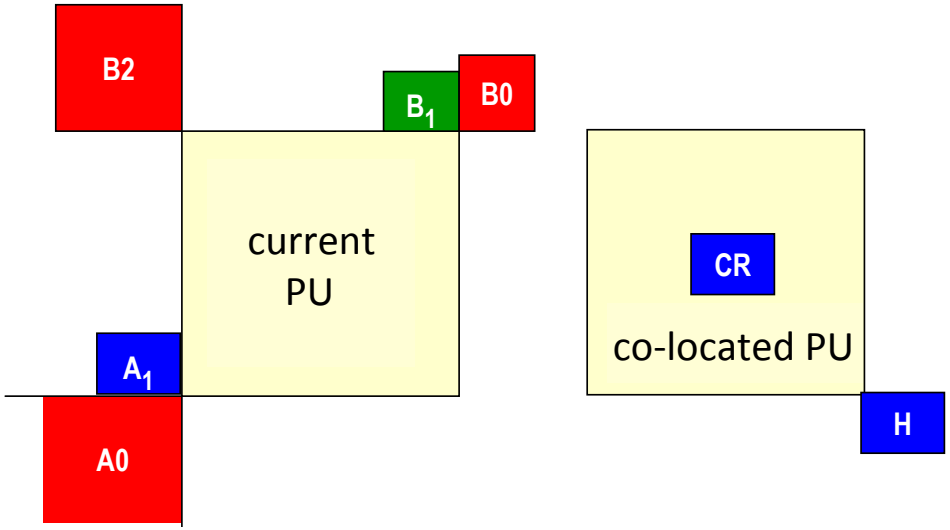
Mode Coding

- Predict modes from neighbors to reduce syntax element bits
  - Intra Prediction Mode



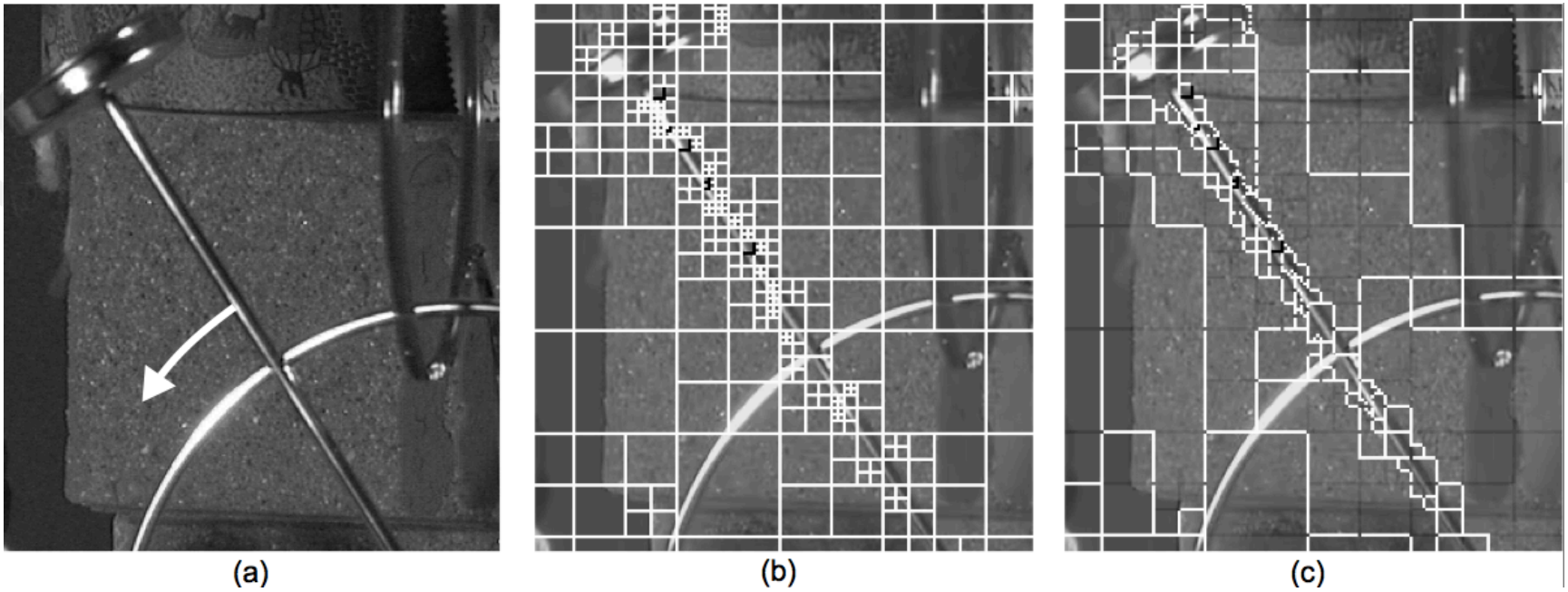
3 candidates

  - Advance Motion Vector Prediction (AMVP), Merge/Skip Mode



2 to 5 candidates

Merge Mode



Moving Object

Without Merge
(many extra motion parameters)

With Merge

B. Bross et al., "Inter Prediction in HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014.

AMVP, Merge, Skip Mode

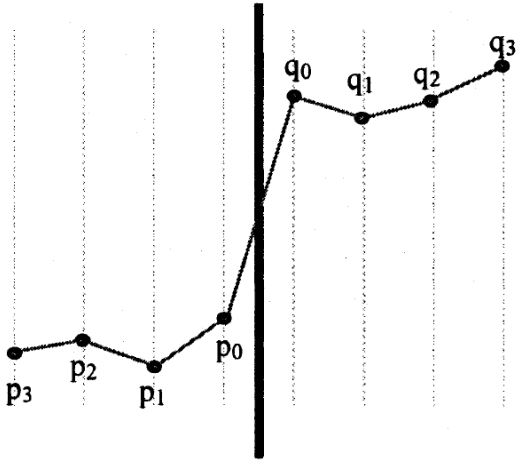
	AMVP	Merge	Skip
Syntax elements	mvp_l0_flag, mvp_l1_flag	merge_flag, merge_idx	cu_skip_flag, merge_idx
Use of neighbors candidates	Predict motion vector	Copy motion data (motion vector, reference index, direction)	Copy motion data (motion vector, reference index, direction); no residual
Number of Candidates	Up to 2	Up to 5 (signaled in slice header)	
Spatial	Up to 2 of 5 (scaling if reference index different)	Up to 4 of 5 (no scaling, only redundancy check)	
Temporal	Up to 1 of 2 (if < 2 spatial candidates)	Up to 1 of 2 (always added to list if available)	
Additional	Zero motion vector (if < 2 spatial or temp candidates)	Bi-predictive candidates and zero motion vector	

In-loop Filtering: Deblocking Filter

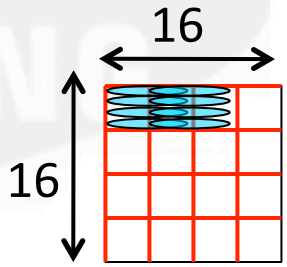
- Removes blocking artifacts due to block based processing
  - Computationally intensive in H.264/AVC

w/o deblocking

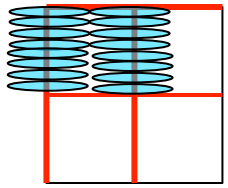
w/ deblocking



- In H.264/AVC, performed on every 4x4 block edge
  - Each macroblock has 128 pixel edges, 32 edge calculations
  - Each 4x4 depends on neighboring 4x4

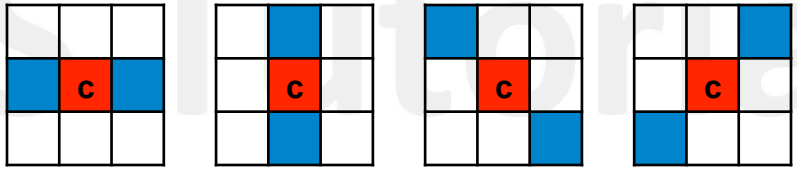


- In HEVC, performed on every 8x8 block edge
  - Each 16x16 CTU has 64 pixel edges, 8 edge calculations
  - All 8x8 are independent (can be processed in parallel)

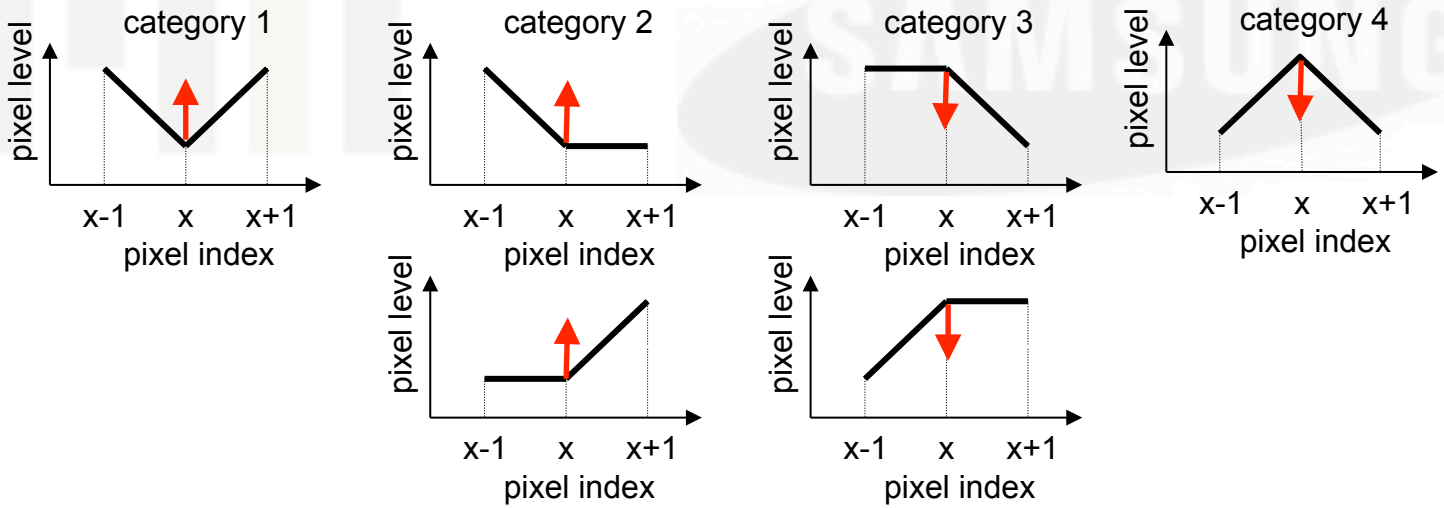


In-loop Filtering: Sample Adaptive Offset (SAO)

- Filter to address local discontinuities
  - Edge Offset and Band Offset
- Check neighbors in one of 4 directions (0, 90, 135, 45 degrees)



- Based on the values of the neighbors, apply one of 4 offsets



In-loop Filtering: Sample Adaptive Offset (SAO)

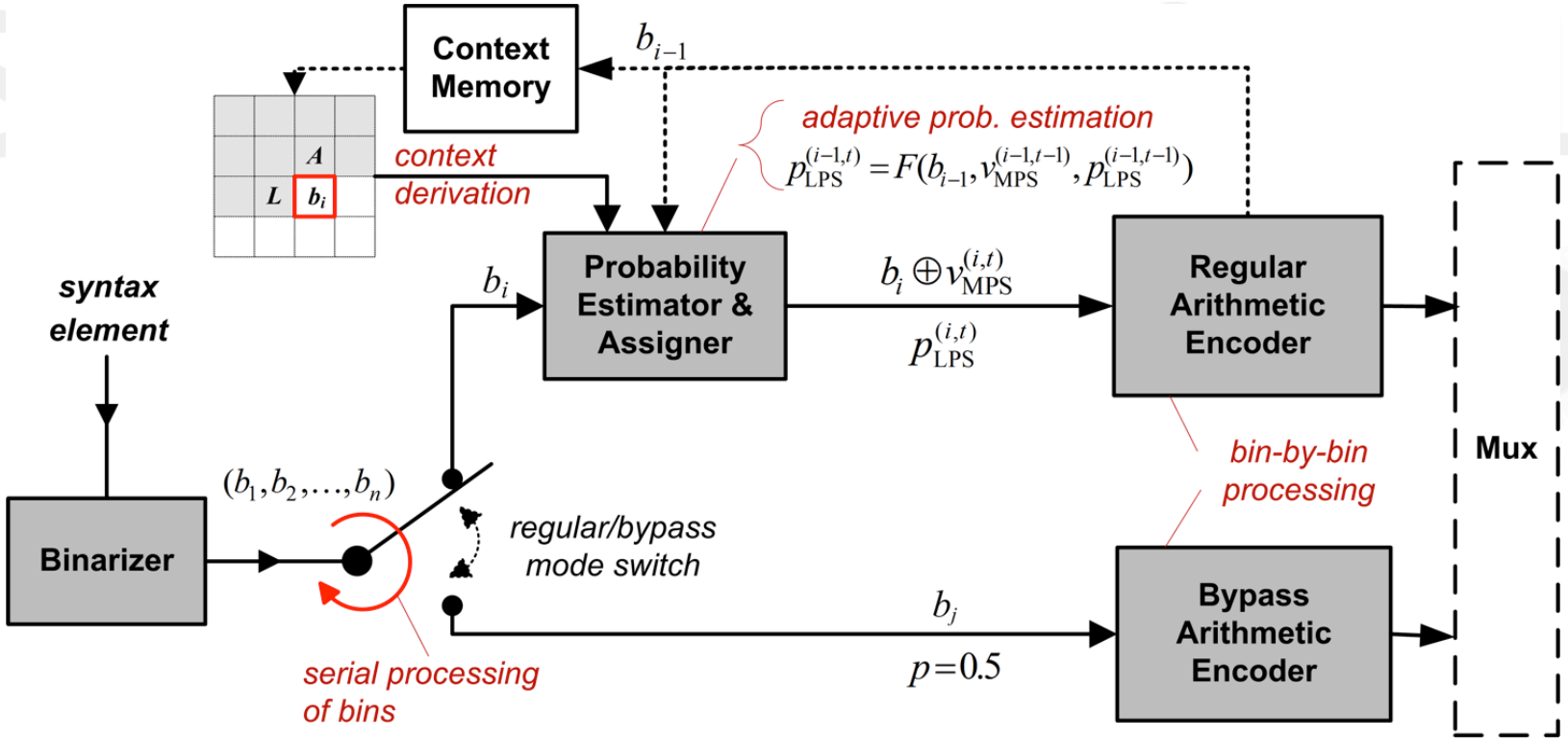
With SAO

Without SAO

C.-M. Fu et al., "Sample Adaptive Offset in the HEVC Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, 2012

Entropy Coding

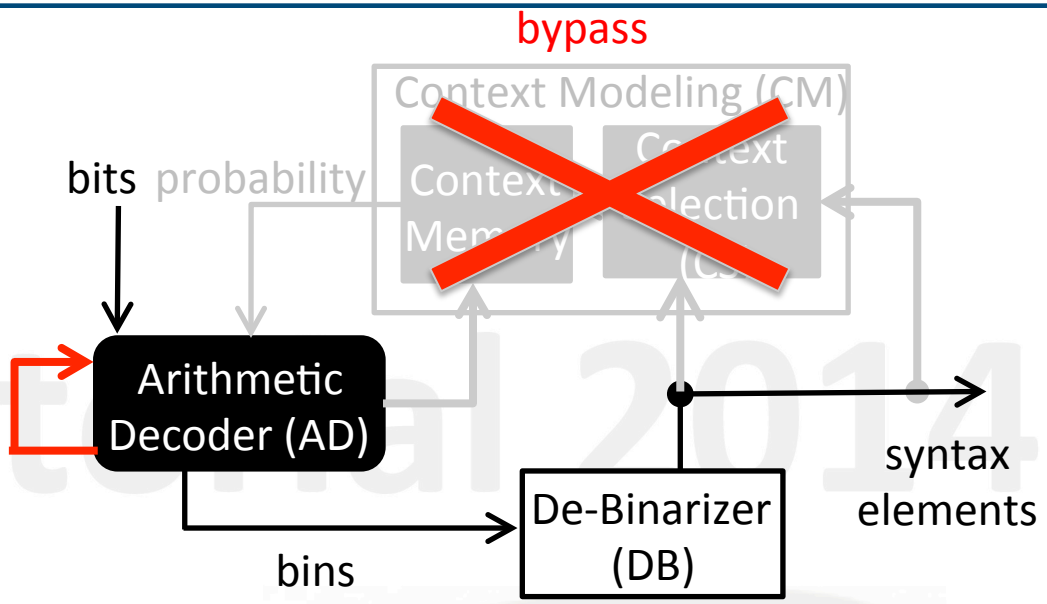
- Lossless compression of syntax elements
- HEVC uses Context Adaptive Binary Arithmetic Coding (CABAC)
  - 10 to 15% higher coding efficiency compared to CAVLC



V. Sze, D. Marpe, "Entropy Coding in HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014.

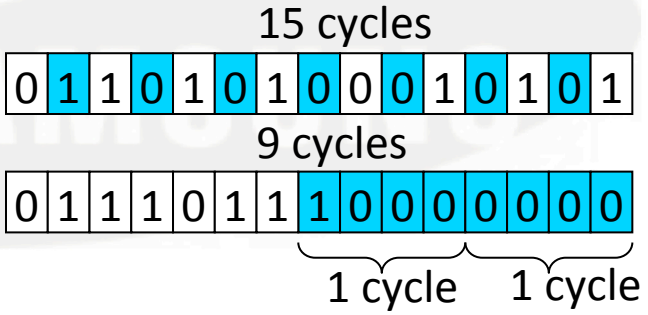
CABAC Throughput Improvements

- Reduce total number of bins
- Reduce context coded bins
- Reduce context dependencies
- **Grouping bypass bins**
- Reduce parsing dependencies
- Reduce memory requirements



Reduction in ***worst case*** bins for 16x16 pixels

	Total bins	Context bins	Bypass bins
H.264/AVC	20861	7805	13056
HEVC	14301	884	13417
Ratio	1.5x	9x	1x

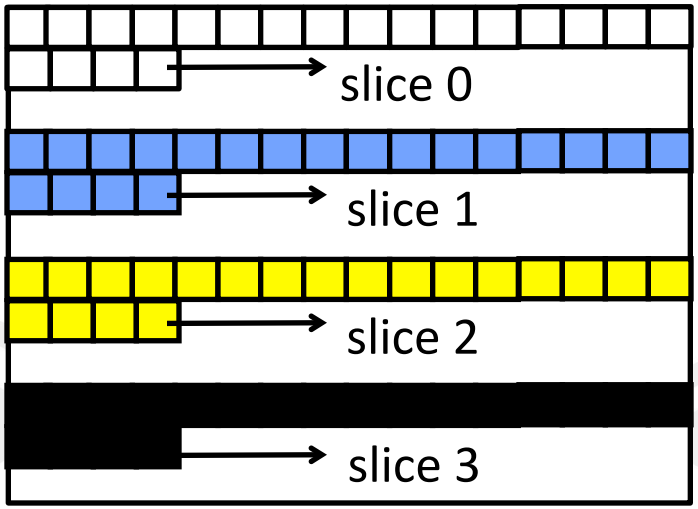


- 3x reduction in context memory
- 20x reduction in line buffer for context selection

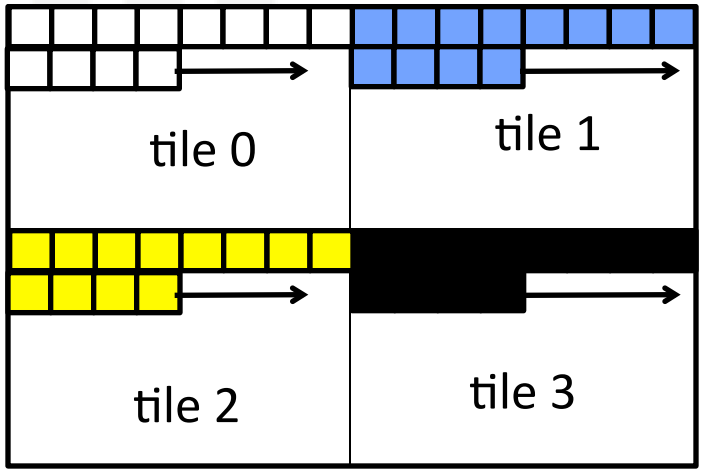
V. Sze, M. Budagavi, "High Throughput CABAC Entropy Coding in HEVC," *IEEE TCSVT*, 2012

High Level Parallel Tools (Multi-Core)

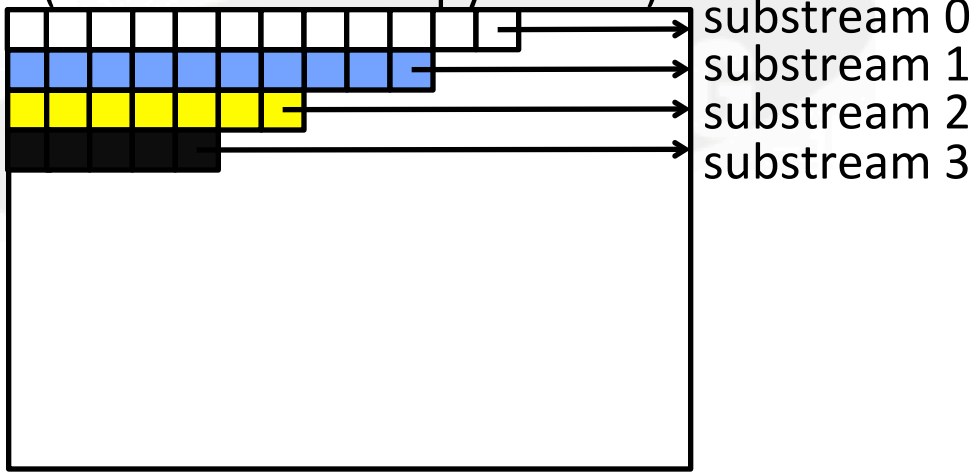
Slices
(also in H.264/AVC)



Tiles



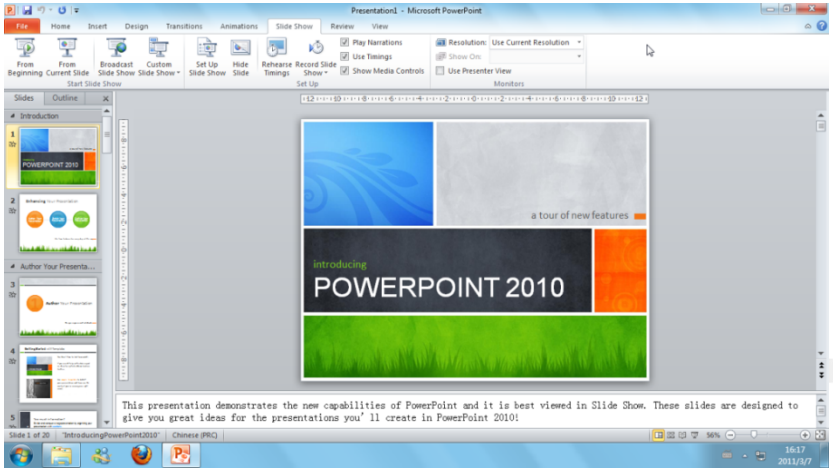
Wavefront Parallel Processing
(Interleaved Entropy Slices*)



*D. Finchelstein, V. Sze, A. P. Chandrakasan, "Multi-core Processing and Efficient On-chip Caching for H.264 and Future Video Decoders," *IEEE Trans. CSVT*, 2009

Additional Modes

- For wireless display and cloud computing, screen content coding should be considered
- Screen content typically has more edges
- Lossless
  - Bypass transform, quantization and in-loop filters
- Transform Skip
  - Bypass transform, but continue to perform quantization and in-loop filters
- I_PCM
  - Signal raw pixels



source: www.techprollc.com

Profiles, Levels, Tiers

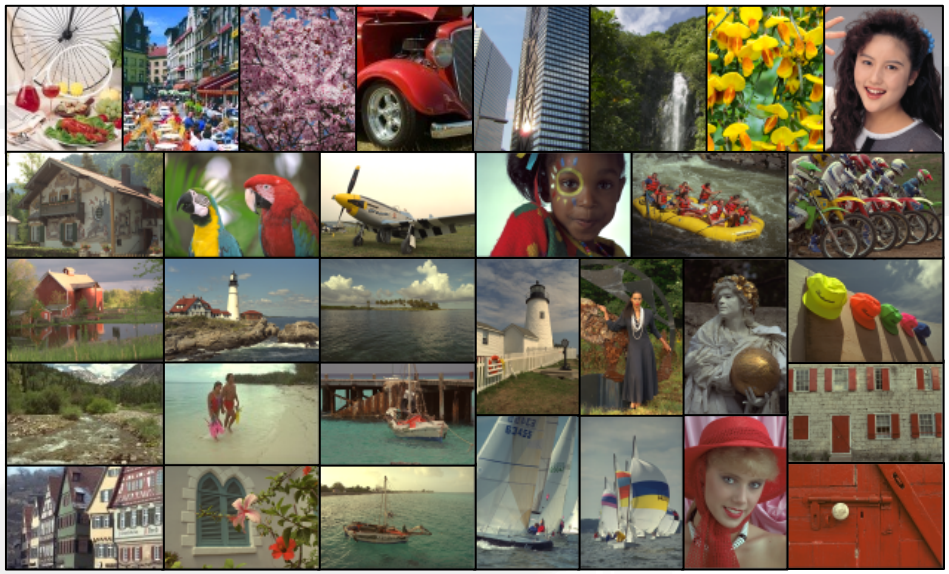
- Profile defines set of tools for different applications
  - Main, Main 10, Main Still Picture
  - 8-bits/sample → 16.78 million colors
  - 10-bits/sample → 1.07 billion colors
- Level defines the maximum supported resolution and frame rate
  - e.g. Level 4.0, 1920x1080 @ 32 fps
  - Level 5.0, 4096x2160 @ 30 fps
- Bit-rates defined by level and tier
  - Main and High (professional)

Level	Max luma sample rate MaxLumaSr (samples/sec)	Max bit rate MaxBR (1000 bits/s)		Min Compression Ratio MinCr
		Main tier	High tier	
1	552 960	128	-	2
2	3 686 400	1 500	-	2
2.1	7 372 800	3 000	-	2
...	...	...	...	...
6	1 069 547 520	60 000	240 000	8
6.1	2 139 095 040	120 000	480 000	8
6.2	4 278 190 080	240 000	800 000	6

Main Still Picture (Intra Coding Only)

- HEVC also provides improved compression for still images

	BD-Rate Reduction
H.264/AVC (intra only)	15.8%
JPEG 2000	22.6%
JPEG XR	30.0%
Web P	31.0%
JPEG	43.0%

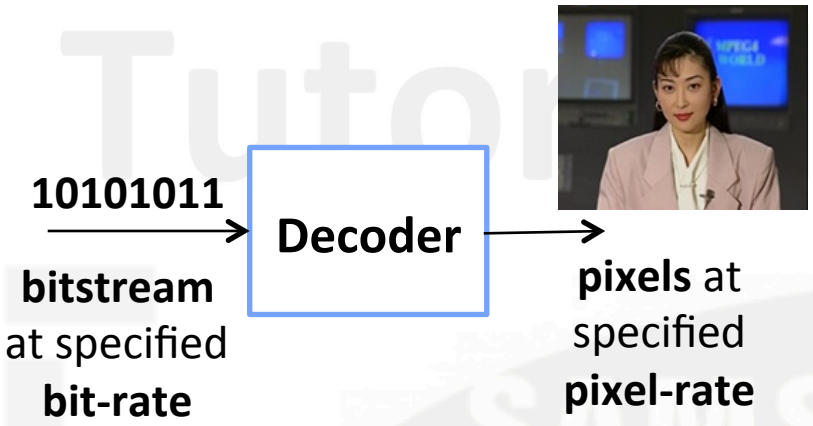


T. Nguyen, D. Marpe, "Performance Comparison of HM 6.0 with Existing Still Image Compression Schemes Using a Test Set of Popular Still Images" JCTVC-I0595, 2012

Part III: Video Codec Implementations

Decoder Design Considerations

- Function
  - Mapping of bitstream to pixels fixed by the standard



- Implementation Requirements
  - *Conformance*: Support all tools for a given profile in the standard
  - *Throughput*: Real-time processing for video playback; **level** specifies pixel-rate and bit-rate

Encoder Design Considerations (1)

- Function
  - Mapping of pixels to standard compliant bitstream
  - Flexibility of selecting which set of encoding tools to use and how to use them (e.g. how to search for best compression mode)

pixels at specified **pixel-rate for real-time applications**

Encoder

10101011

bitstream at specified **bit-rate or compression ratio**

Encoder Design Considerations (2)

- Implementation Requirements
  - *Conformance*: Must generate a bitstream that is decodable by a standard compliant decoder (for a given profile)
  - *Throughput*: For real-time applications, need to meet pixel-rate requirements; can be done off-line for storage applications
  - *Bit-rate/Compression Ratio:* For given application, must meet minimum compression requirements
  - *Compression ratio vs. Complexity:* Find compression mode that meets compression requirements under complexity constraint

Decoder design requires architecture innovations, while encoder design requires both algorithm and architecture innovations

Multimedia Platforms

	Desktop CPU [1]	Mobile CPU [1]	GPU+CPU [2]	DSP [3]	FPGA [4]	ASIC [5,6]
Flexibility	High	High	Med/High	Med	Med	Low
Development Cost	Low	Low	Low/Med	Med	Med	High
Speed/ Throughput	Low/Med	Low	Med	Med	Med	High
Power Consumption	High	Med	High	Med	Med	Low

Examples of HEVC implementations

[1] F. Bossen et al., "HEVC Complexity and Implementation Analysis," *IEEE TCSVT,* 2012

[2] Ittanim Systems, "Compute accelerated HEVC decoder on ARM® MaliTM-T600 GPUs"

[3] F. Pescador et al., "On an implementation of HEVC video decoders with DSP technology," *IEEE ICCE,* 2013

[4] S. Cho, H. Kim, "Implementation of a HEVC Hardware Decoder," JCTVC-L0098, 2013

[5] C.-T. Huang et al. "A 249Mpixel/s HEVC video-decoder chip for Quad Full HD applications," *IEEE ISSCC*, 2013.

[6] S.-F. Tsai et al. "A 1062Mpixels/s 8192× 4320p High Efficiency Video Coding (H.265) encoder chip," *IEEE VLSIC*, 2013.

Implementation Requirements

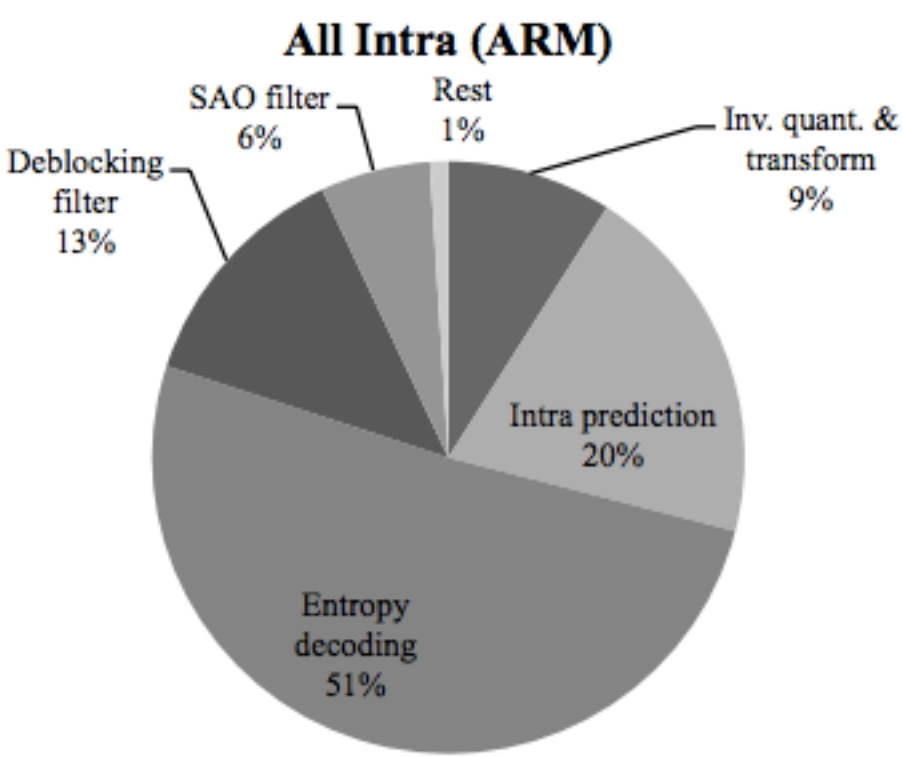
- Throughput
  - Achieve target pixel-rate and bit-rate for real-time applications
  - Reduce latency of bits to pixels and pixels to bits for interactive applications
  - Techniques: parallelism, pipelining, eliminate stalls
- Energy and Power Consumption
  - Minimize energy consumption to extend battery life for portable devices
  - Minimize power consumption to reduce heat dissipation
  - Techniques: voltage scaling, frequency scaling, power gating, number of ops
- Platform Cost
  - Reduce amount of data to be stored in memory and amount of logic (e.g. gates in ASIC, number of cores for processors) to reduce size of chip
  - Reduce bandwidth requirements such as reads/writes from memory to reduce demands on off-chip components
  - Techniques: shared computations, on-the-fly processing, caching

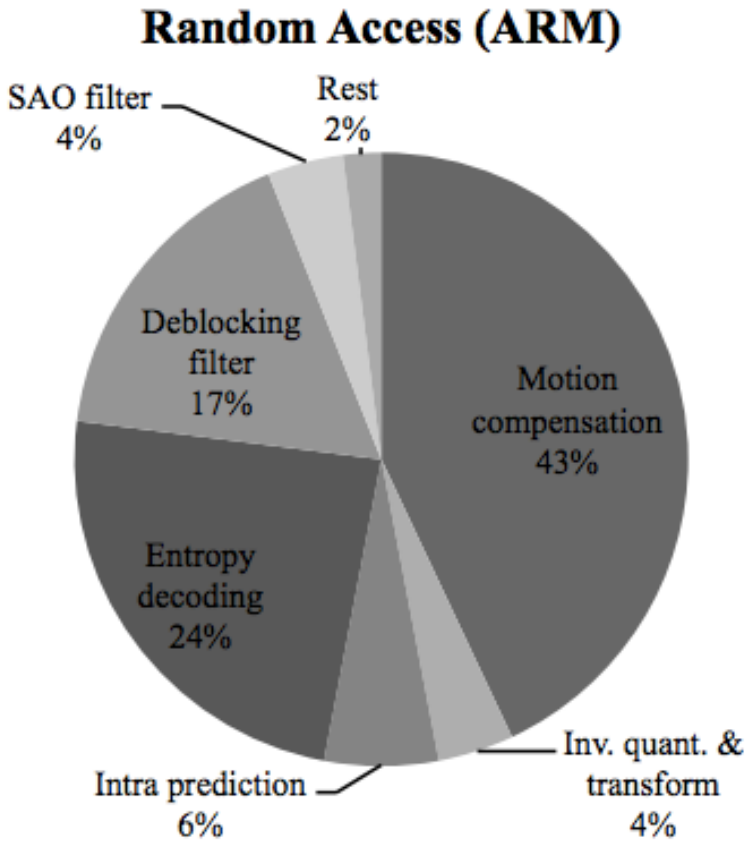
Software HEVC Decoder

- ARMv7 1.3GHz (mobile processor) [Bossen, JCTVC-K0327, 2012]
  - Dual core, but decoding on single thread (other thread for display)
  - 1080p @ 24 fps at 2Mbps (16 picture buffer to average workload)
- Intel i7 Core 2.6 GHz (desktop processor) [Bossen et al., TCSVT, 2012]
  - Single core, single thread
  - 1080p @ 60 fps at 7Mbps
- Multi-thread Intel Core i7 2.7 GHz [Suzuki et al., JCTVC-L0098, 2013]
  - 4 cores / 4 threads (parallel GOPs)
  - 3840x2160 @ 76 fps at 12Mbps [cropped 8K content]
- Multi-thread Intel X5680 3.3 GHz [Chi et al., TCSVT, 2012]
  - 2x6 cores/12 threads (parallel Tiles, WPP)
  - 3840x2160 @ 24 fps at ~12Mbps (QP=37)
  - 3840x2160 @ 14 fps at ~170Mbps (QP=22)

Software HEVC Decoder

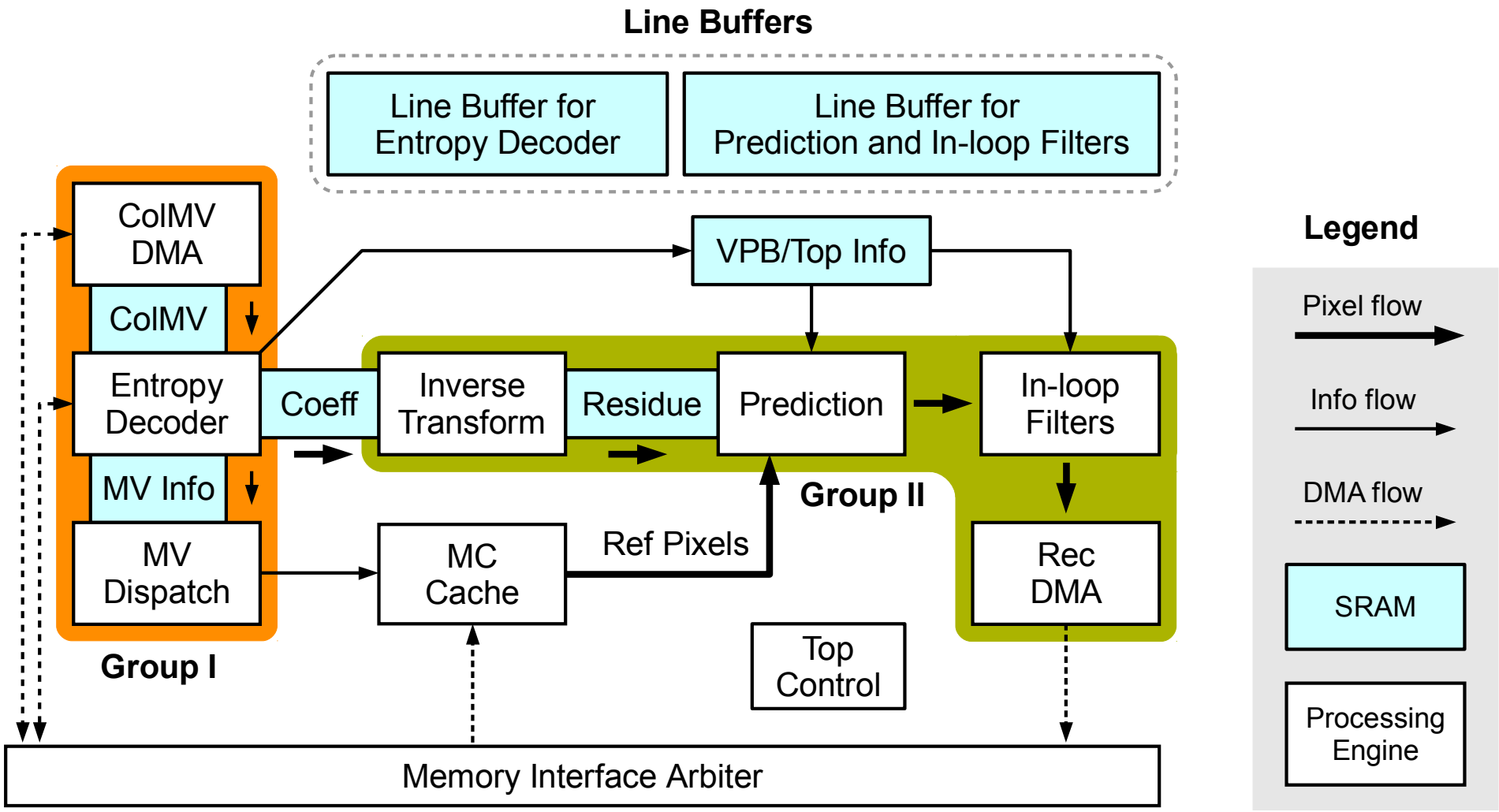
Workload for different modules





F. Bossen et al., "HEVC Complexity and Implementation Analysis," *IEEE Transactions on Circuits and Systems for Video Technology,* 2012

Hardware HEVC Decoder Architecture

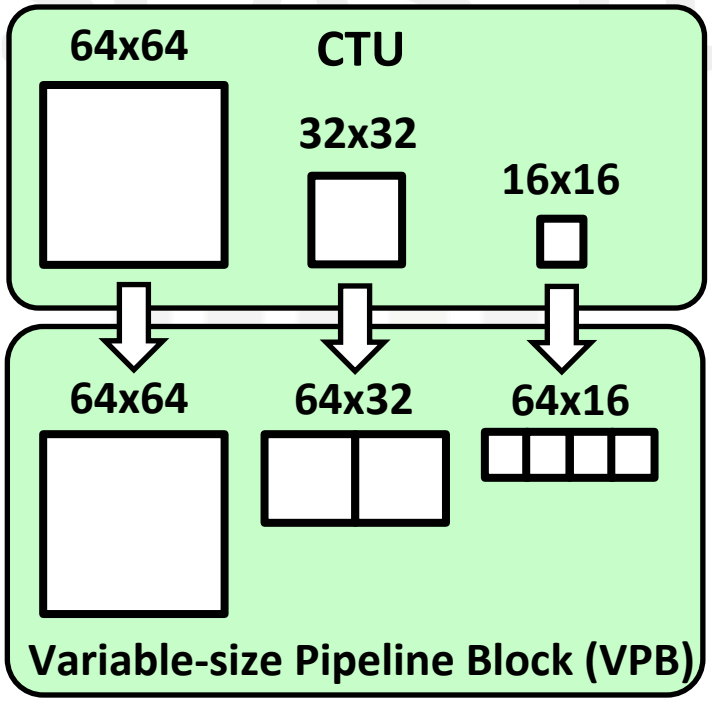


M. Tikekar et al., "Decoder Hardware Architecture for HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014.

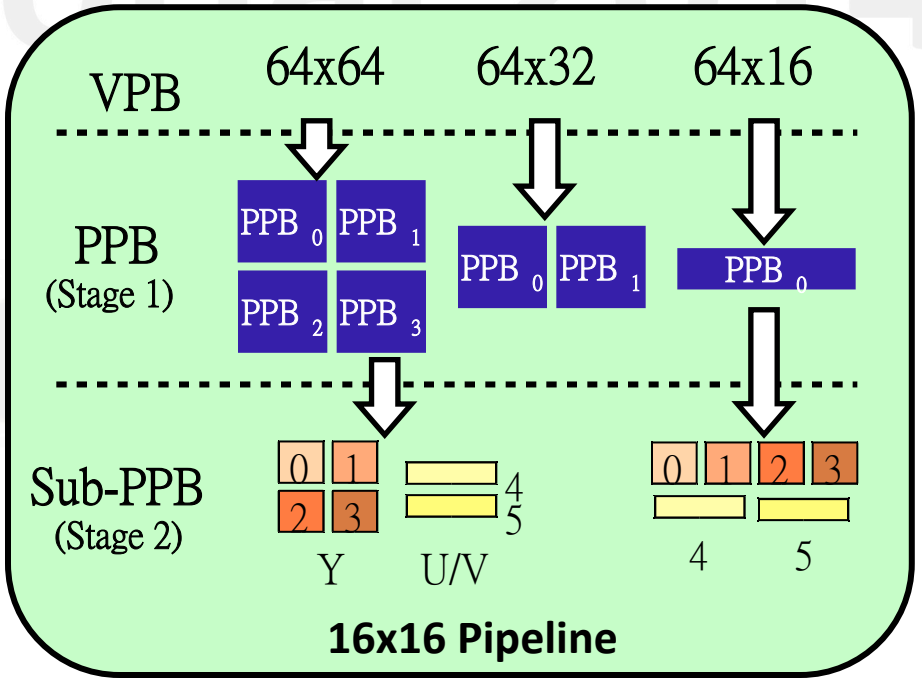
Pipelining HEVC Decoder

- Variable-size pipelining to support a diverse set of CTU, CU, and PU sizes (select size to balance memory cost vs. data reuse)

System level pipeline
(between Inv. Transform, Prediction and In-Loop Filters)



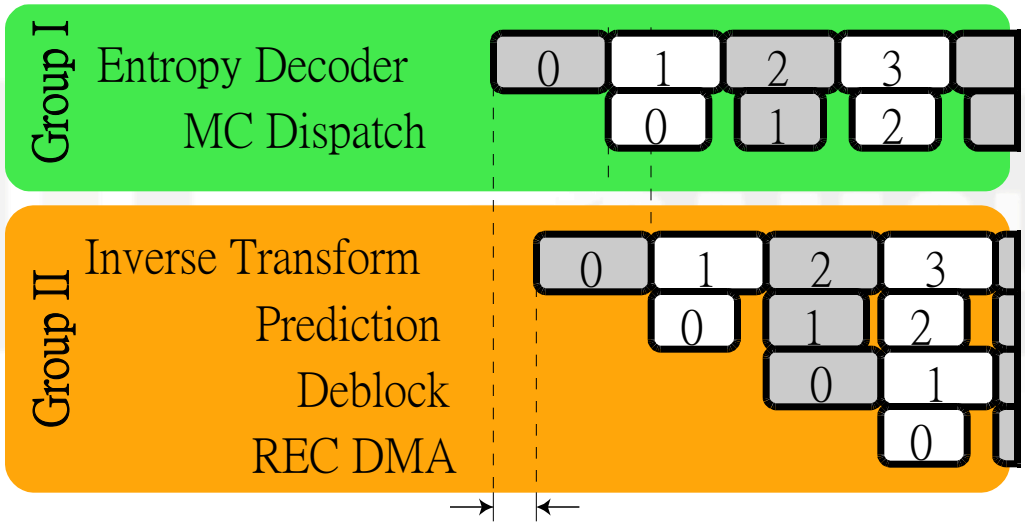
Prediction level pipeline
(within Prediction module)



Source: C.-T. Huang et al., "A 249Mpixels/s HEVC Video Decoder Chip for Quad Full HD Applications," *IEEE ISSCC*, 2013.

Decoupling Entropy Coding

- Workload of entropy decoding based on bit-rate (bin-rate), while rest of decoder depends on pixel-rate
- Use FIFO to absorb variations in workload
  - Higher FIFO depth results in less stalls due to averaging, but longer latency and higher memory cost

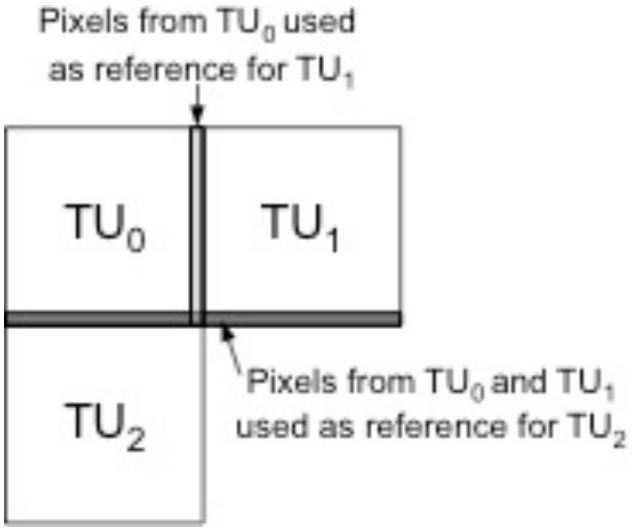


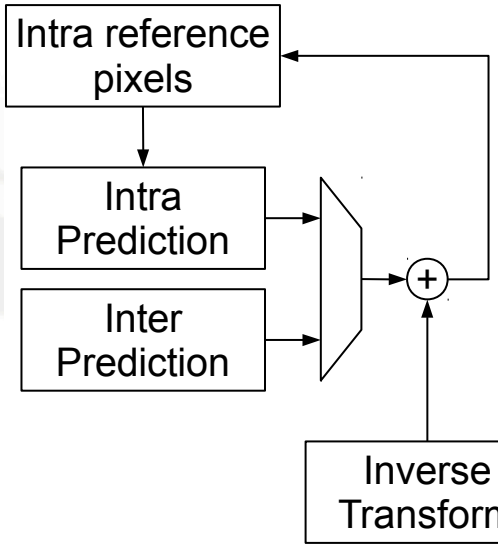
Coefficients in TU FIFO

Source: C.-T. Huang et al., "A 249Mpixels/s HEVC Video Decoder Chip for Quad Full HD Applications," *IEEE ISSCC*, 2013.

Intra Prediction

- Reference sample processing
  - Reference pixel buffer to store neighboring pixels (padding when not available)
  - Apply smoothing filter on pixels depending on mode
- Feedback loop at TU granularity
  - Update reference pixel buffer accordingly



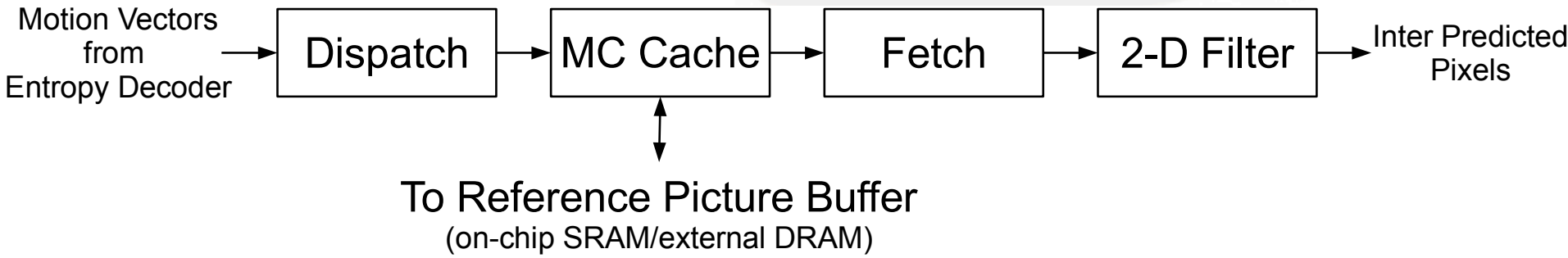


TU granularity feedback

M. Tikekar et al., "Decoder Hardware Architecture for HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014.

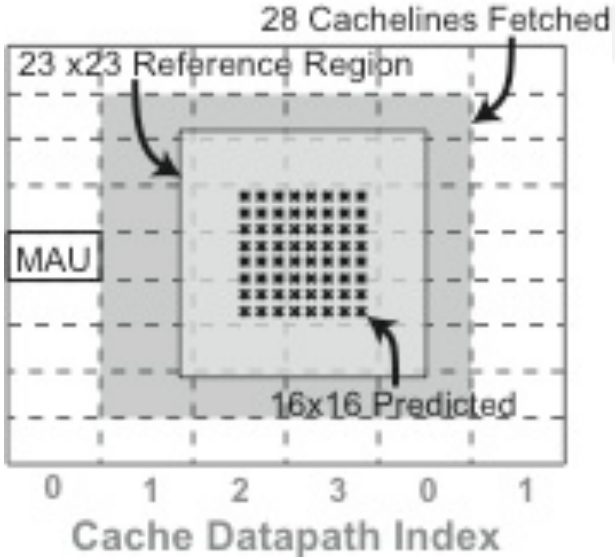
Inter Prediction

- Read samples from reference picture (typically stored in off-chip picture buffer)
  - Use cache to reduce off-chip memory bandwidth
- Interpolation pixels used a 2-D separable filter for fractional motion vectors
  - Multiple pixels can be interpolated in parallel (share input pixels)
- Smaller blocks have larger read overhead (for fractional mv)
  - NxN requires (N+7)x(N+7) pixel reads → 4x4 inter-PU not supported in HEVC



MC Cache and Picture Buffer

- Minimize redundant reads from off-chip memory (DRAM)
- MC Cache design considerations
  - Sufficient throughput to support worst case PU
  - Detect redundant reads and handle latency of DRAM
- Store pixels in DRAM to minimize row changes (cycle overhead)
  - Avoid reading two rows from same bank for a given reference region



0	1	4	5
2	3	6	7
4	5	0	1
6	7	2	3

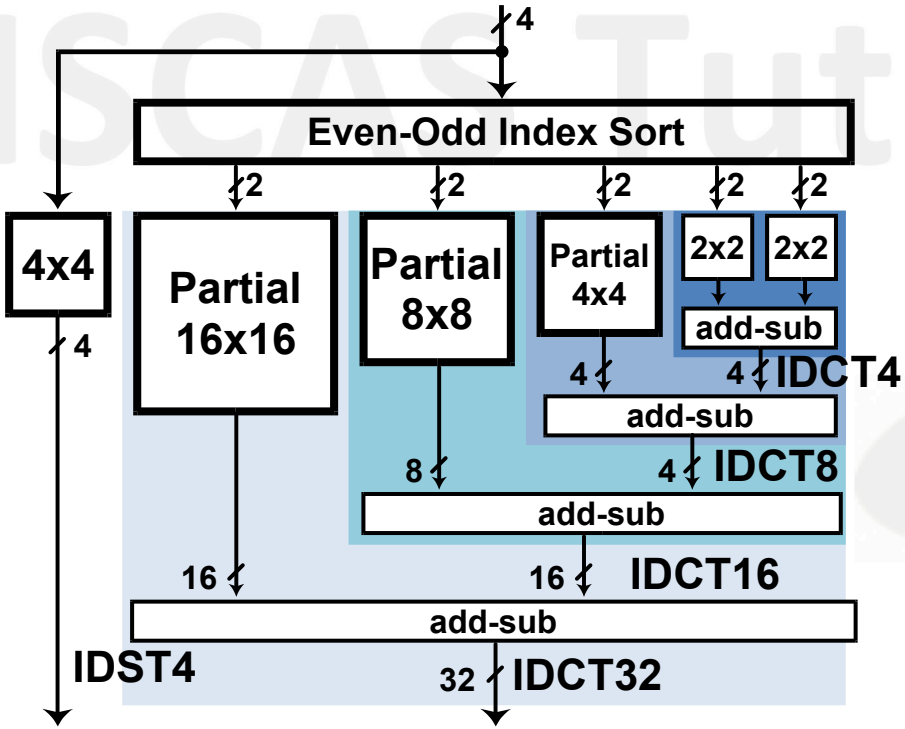
20% reduction in overhead cycles

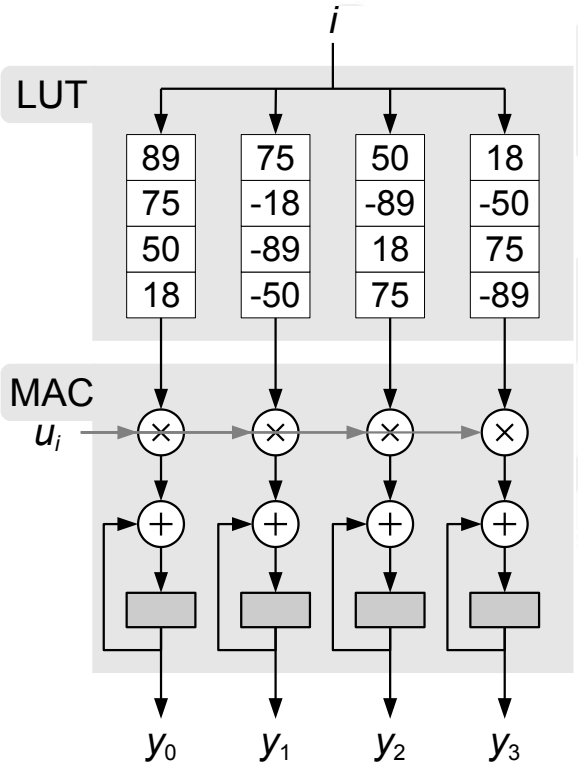
 # = bank in DRAM

M. Tikekar et al., "Decoder Hardware Architecture for HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014.

Inverse Transform

- Larger transform → More computation
  - Share coefficients across transform sizes and within transform to reduce area cost





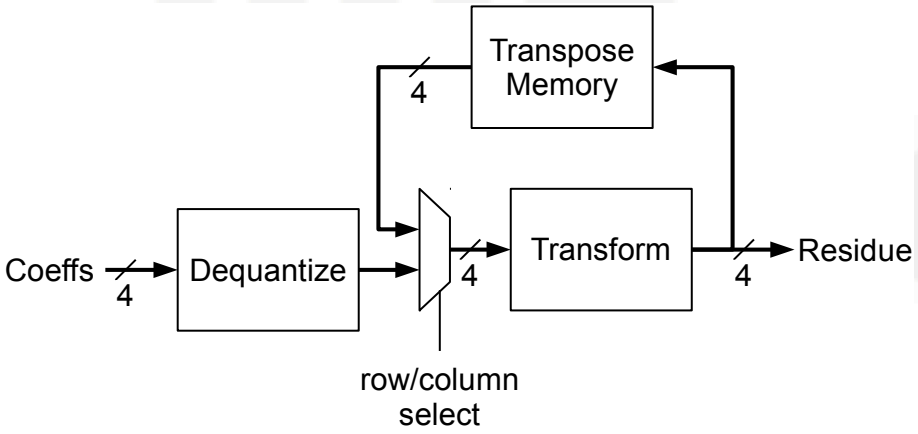
30% reduction in area cost

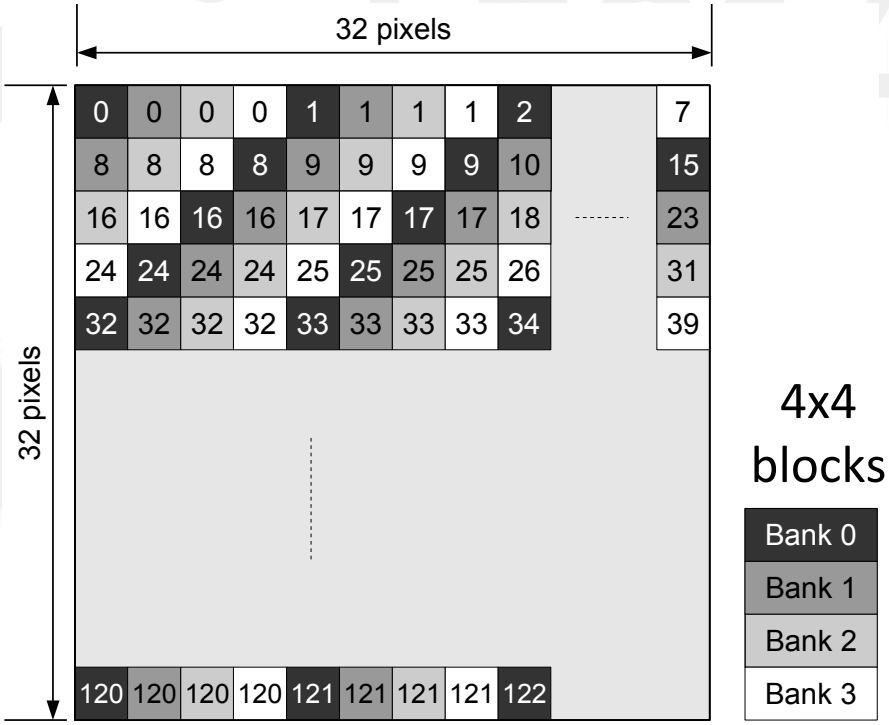
M. Tikekar et al., "Decoder Hardware Architecture for HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014.

Inverse Transform

- Larger transform → Larger transpose memory
  - Use SRAM rather than registers to reduce area cost
  - SRAM has limited read/write ports (requires careful mapping)

4 pixels/cycle throughput per 1-D transform

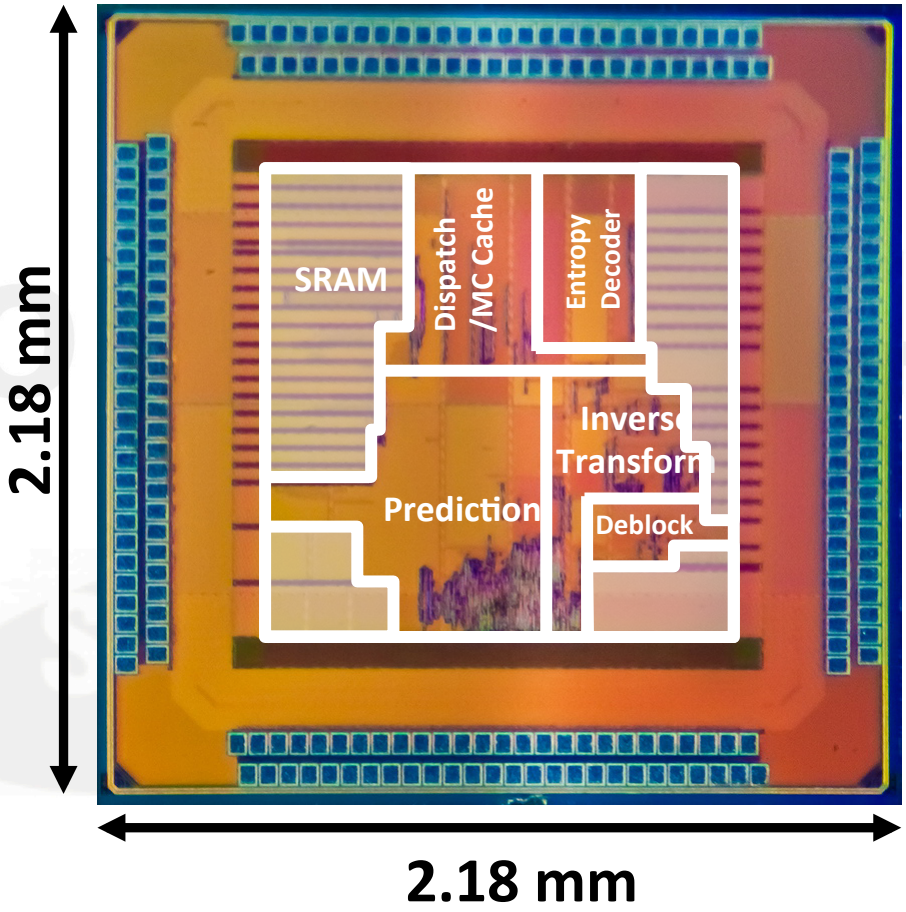




M. Tikekar et al., "Decoder Hardware Architecture for HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures,* Springer, 2014.

Hardware HEVC Decoder

Video Coding Standard	HEVC (HM4)
Technology	TSMC 40-nm
Core Area	1.33 x 1.33 mm
Gate Count	715k
On-Chip Memory (SRAM)	124 kB
Resolution / Frame Rate	4kx2k @ 30fps (3840x2160)
Frequency	200 MHz
Core Voltage	0.9 V
Power	76 mW

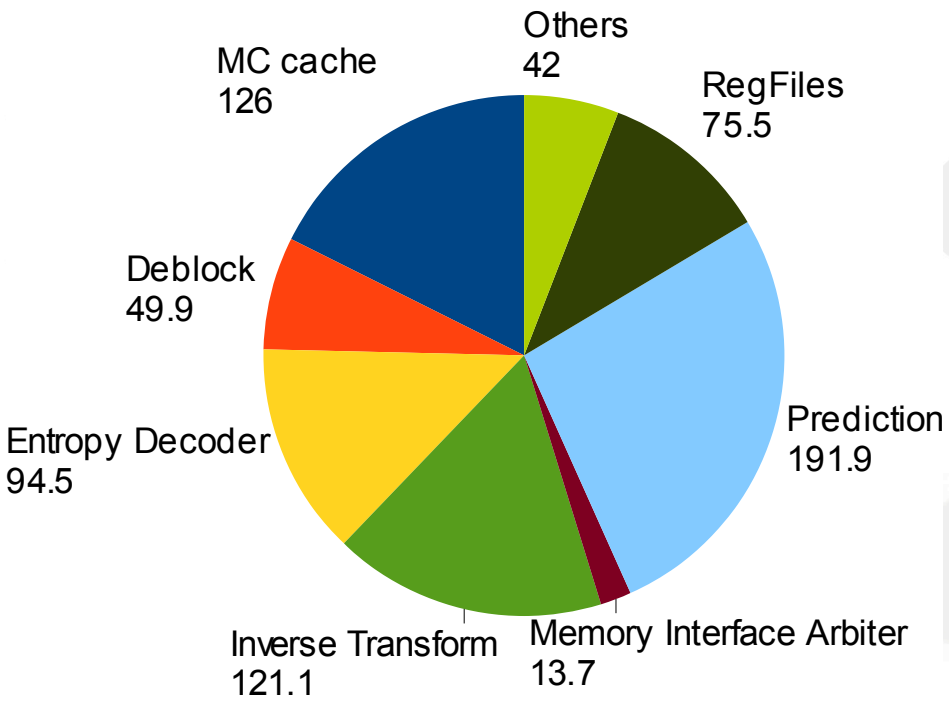


C.-T. Huang et al., "A 249Mpixels/s HEVC Video Decoder Chip for Quad Full HD Applications," *IEEE ISSCC*, 2013

Area Breakdown

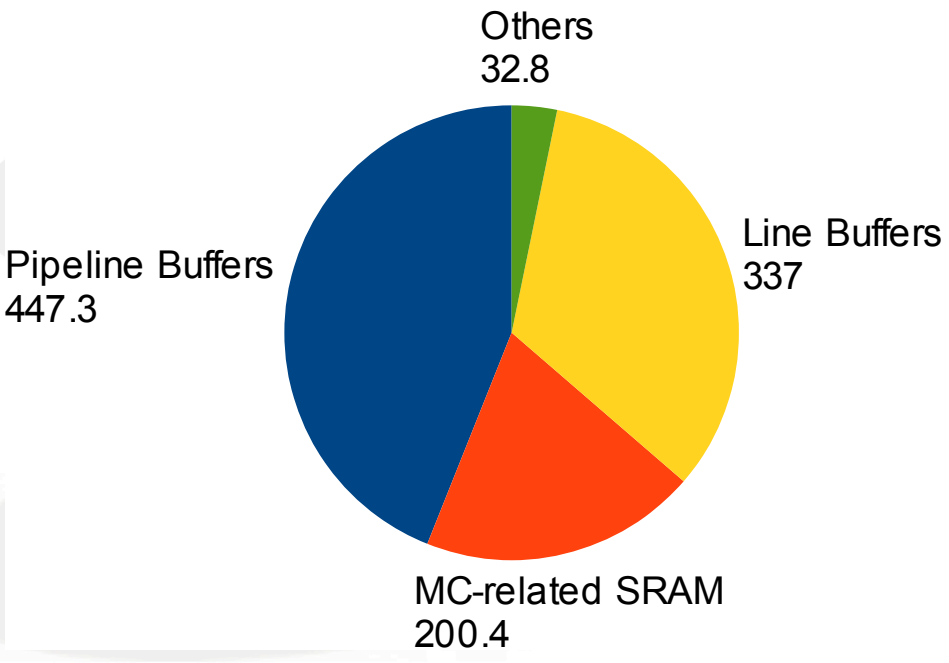
Logic

[kgates]



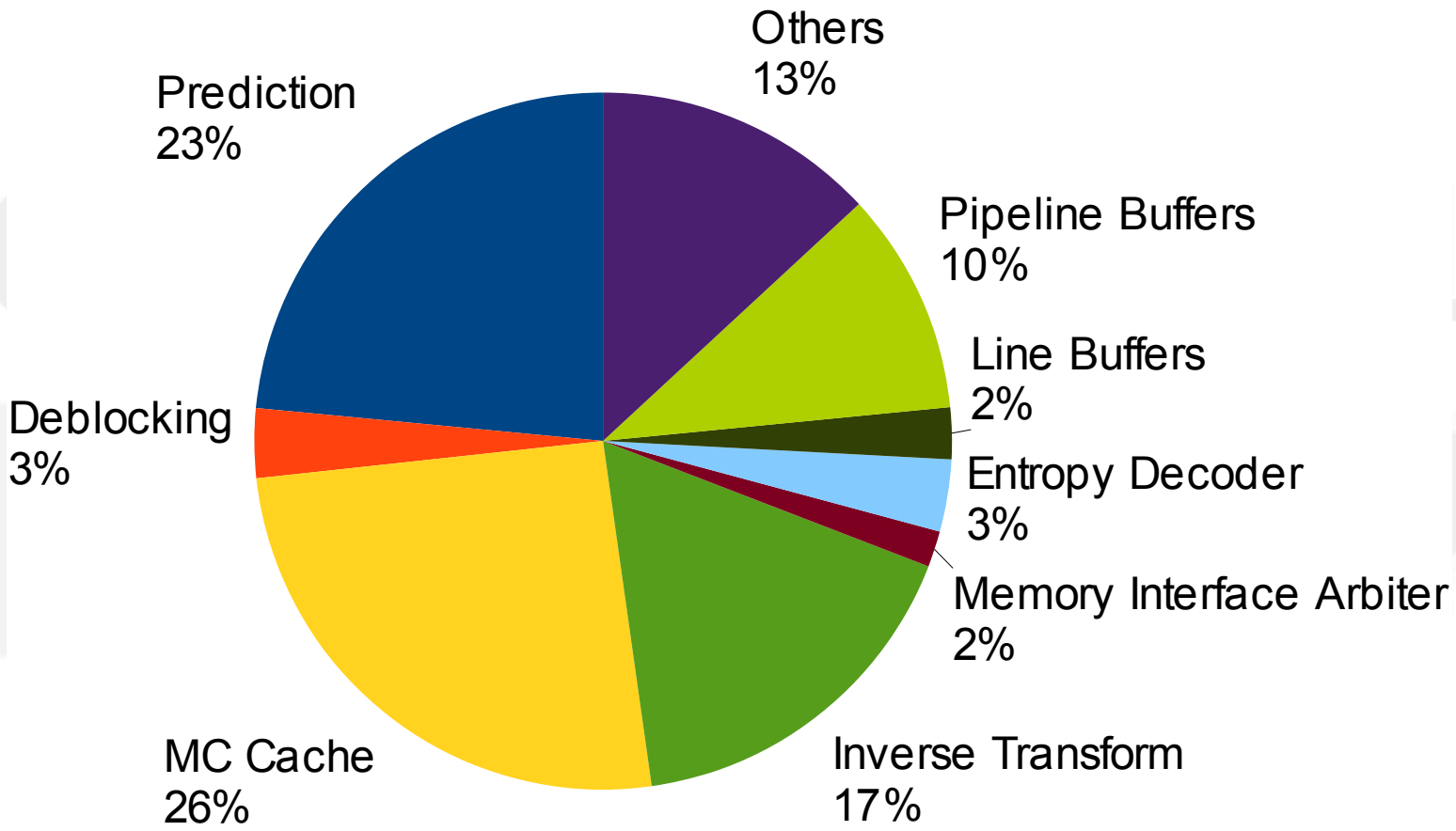
Memory (SRAM)

[kbits]



M. Tikekar et al., "Decoder Hardware Architecture for HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014.

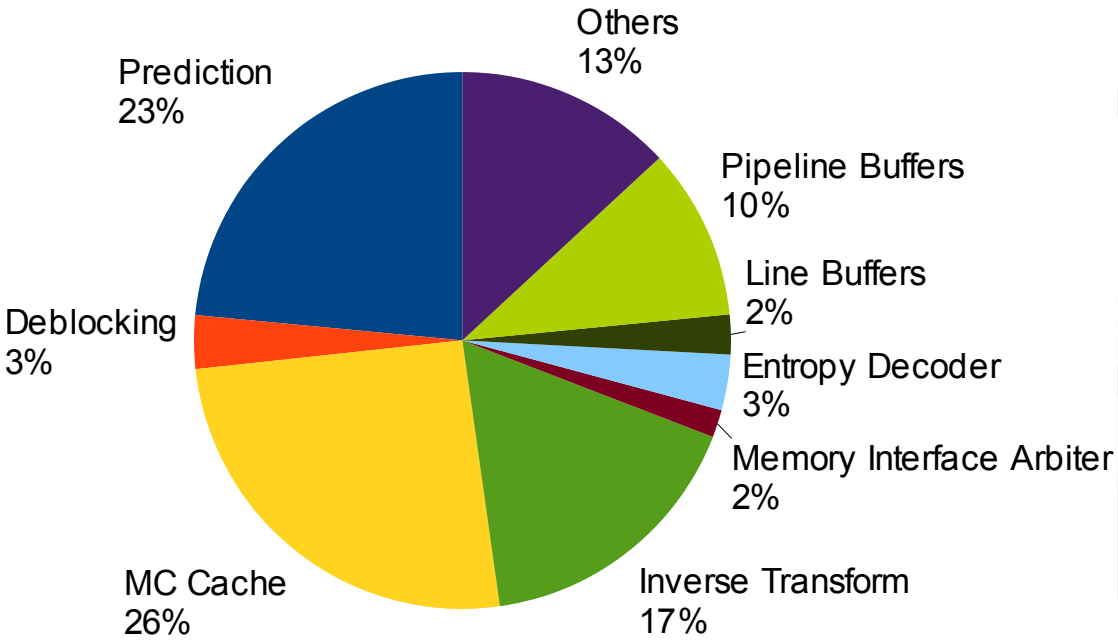
Power Breakdown



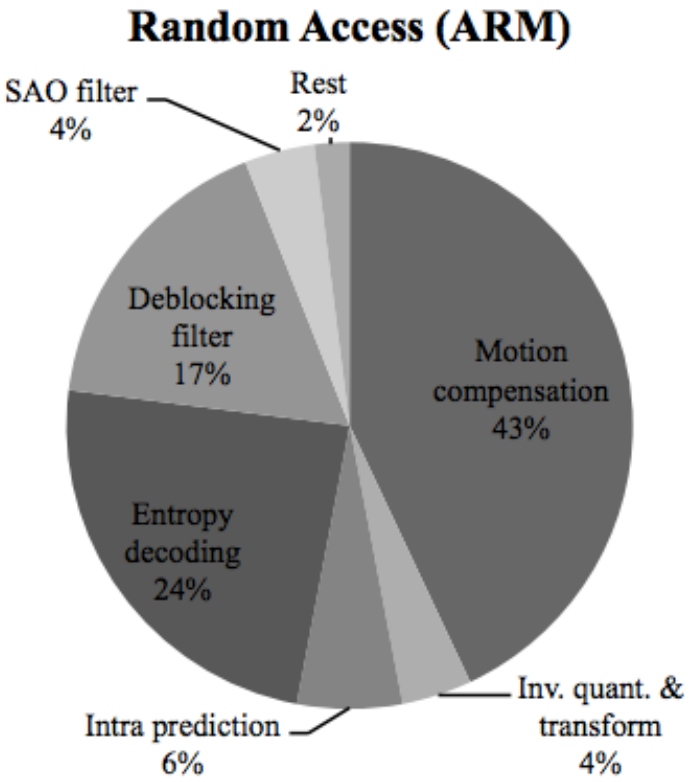
M. Tikekar et al., “Decoder Hardware Architecture for HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014.

Hardware vs. Software

Hardware (power)



Software (cycles)



ASIC Decoder Comparison

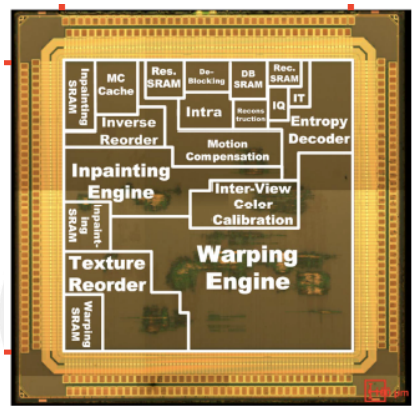
	This Work	ISSCC'12 [2]	ISSCC'10 [3]	ISSCC'06 [4]
Standard	**HEVC ("H.265") WD4**	H.264/AVC HP/MVC	H.264/AVC HP/SVC/MVC	H.264/AVC MP
Max Specification	**3840x2160 @30fps**	7680x4320 @60fps	4096x2160 @24fps	1920x1080 @30fps
Gate Count	**715K**	1338K	414K	160K
On-Chip SRAM	**124KB**	80KB	9KB	5KB
Technology	**40nm/0.9V**	65nm/1.2V	90nm/1.0V	0.18μm/1.8V
Normalized Core Power*	**0.31nJ/pixel**	0.21nJ/pixel	0.28nJ/pixel	5.11nJ/pixel
Normalized DRAM Power*	**0.88nJ/pixel****	1.27nJ/pixel	N/A	N/A
Normalized System Power***	**1.19nJ/pixel**	1.48nJ/pixel	N/A	N/A
DRAM Configuration	**32b DDR3**	64b DDR2	N/A	32b DDR + 32b SDR

* **Power for max specification**
** **Modeled by [5]**
*** **System Power = Core Power + DRAM Power**

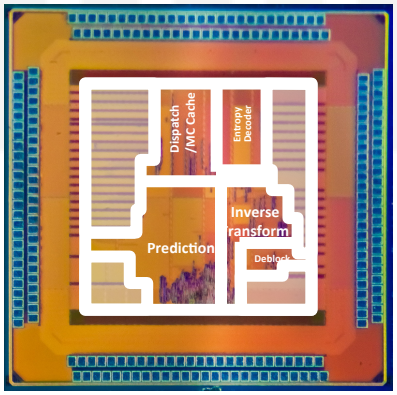
Slide Source: C.-T. Huang et al., "A 249Mpixels/s HEVC Video Decoder Chip for Quad Full HD Applications," *IEEE ISSCC*, 2013.

Decoder Power Comparison

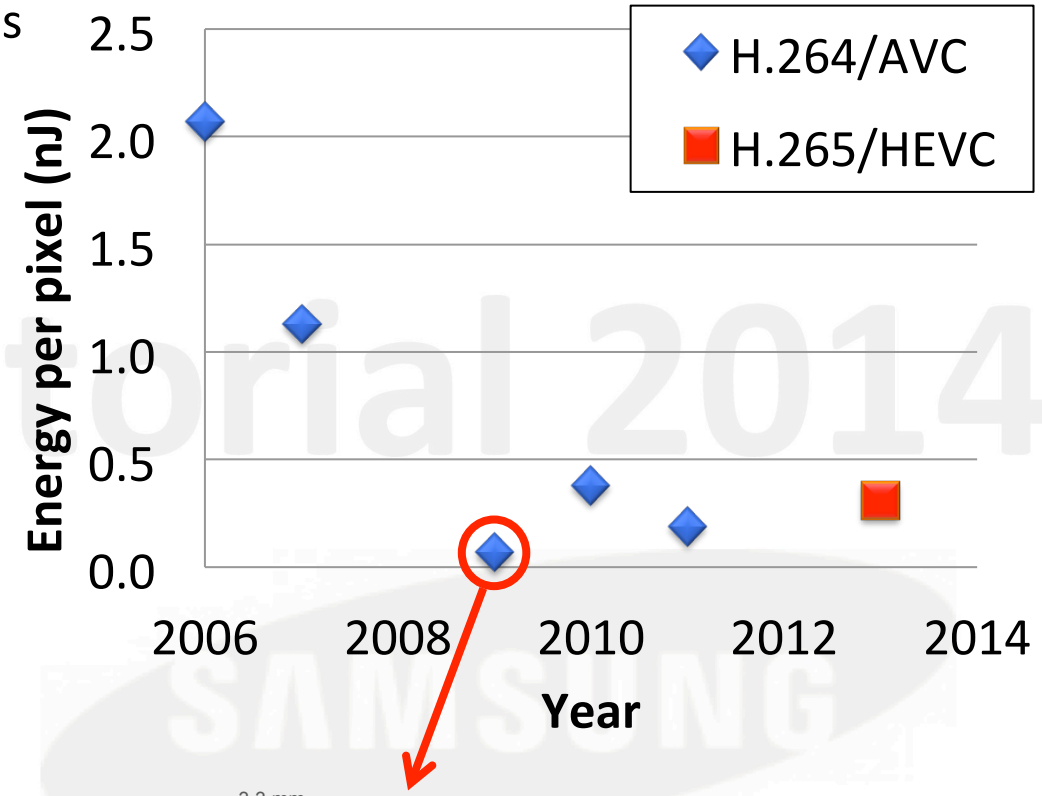
TSMC 40nm, 0.9V Ultra-HD 4K @ 30 fps

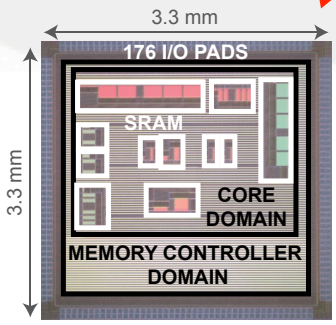


H.264/AVC Decoder (51mW)
P.K. Tsung et al. (NTU), ISSCC 2011



H.265/HEVC [WD4] Decoder (76mW)
C.T. Huang et al. (MIT), ISSCC 2013

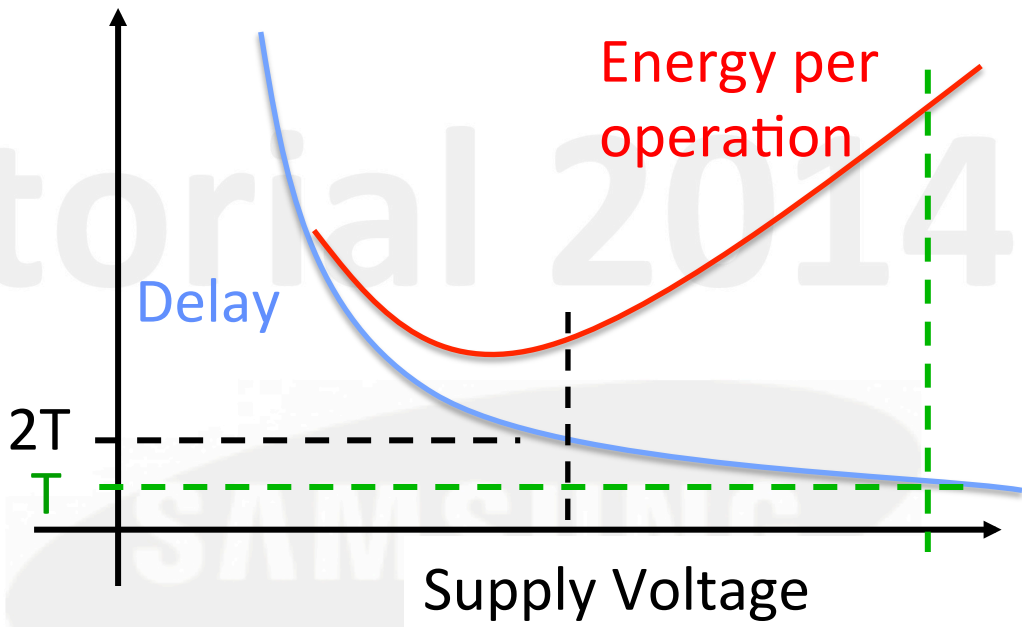




0.7-V 720p-HD @ 30 fps
H.264/AVC Decoder (2mW)
Sze et al. (MIT), JSSC 2009

Low Power Approaches

- Operate at voltage near minimum energy point
- Utilize parallelism and pipelining to achieve performance
- Adaptive/Dynamic voltage frequency scaling
- Optimize access patterns to reduce memory power



Reduce Cycles → Reduce Freq. → Reduce Voltage → Reduce Power

V. Sze et al., "A 0.7-V 1.8-mW H.264/AVC 720p Video Decoder," *IEEE Journal of Solid State Circuits*, 2009.

Encoder Decisions

- Encoder must search for mode that gives the "best" compression. Some of the key decisions include
  - CU and PU size
  - Inter or Intra CU
  - Motion Vector
  - Intra Prediction Mode
- "Best" compression is defined using a rate-distortion cost

$$D + \lambda \cdot R$$

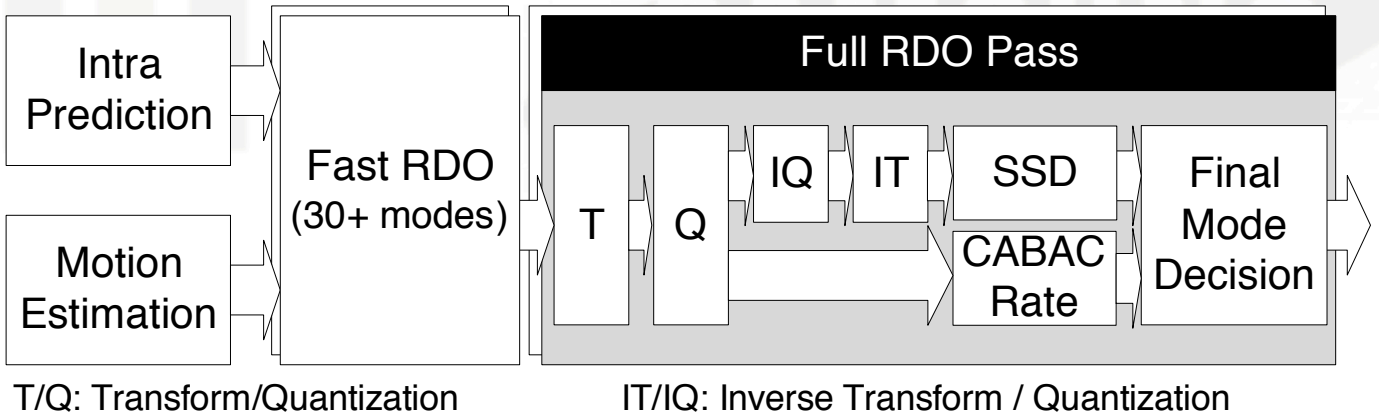
Perform rate-distortion optimization (RDO)

- where
  - D is the distortion between the original and the compressed image (a measure of the visual quality of the compression)
  - R is a measure of the number of bits required to signal the compressed image
  - λ is the Lagrangian multiplier that weights the distortion and rate costs

Full vs. Fast RDO

- Full RDO
  - Distortion based on sum of squared differences (SSD), includes quantization
  - Rate based on entropy coded bits of prediction info and quantized coefficients
- Fast RDO
  - Distortion approximation based on sum of absolute differences (SAD) or sum of absolute transformed differences (SATD)
  - Rate approximation based on prediction info bits (intra mode or motion vector); Can include number of non-zero coefficients to predict coefficient bits

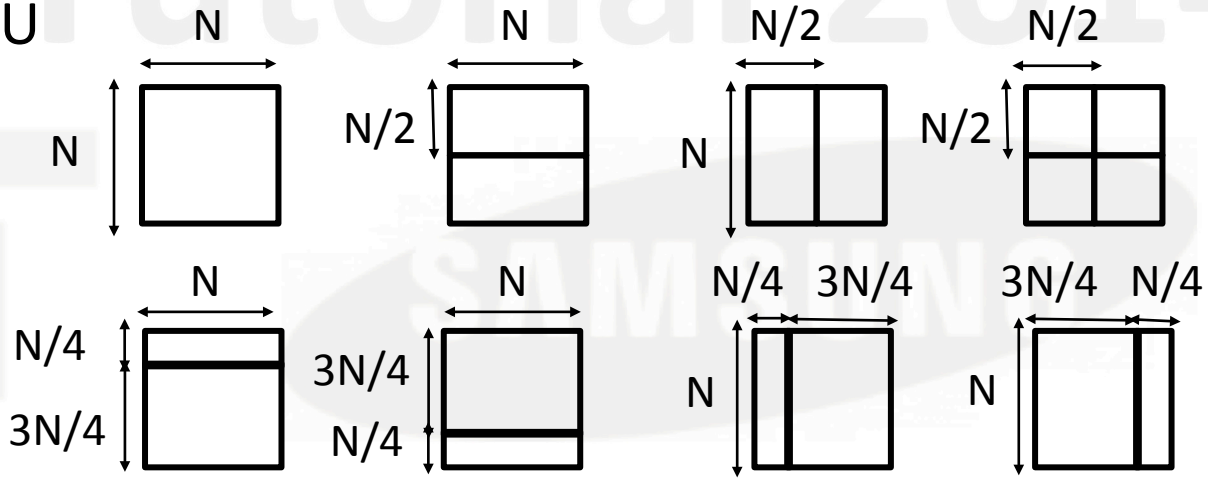
RDO Flow in HM



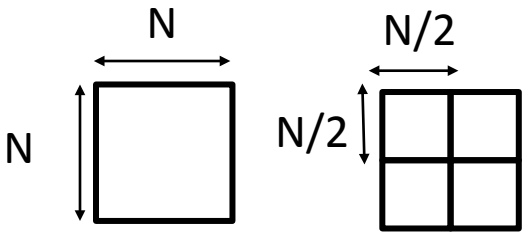
S. -F. Tsai et al., "Encoder Hardware Architecture for HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014.

CU and PU decisions

- The encoder must decide to how best divide a CTU into CU, and how to divide the CUs into PUs **(based on full RDO in HM)**
- For CTU of 64x64
  - CU options: 64x64, 32x32, 16x16, 8x8
- For Inter-coded CU
  - PU options

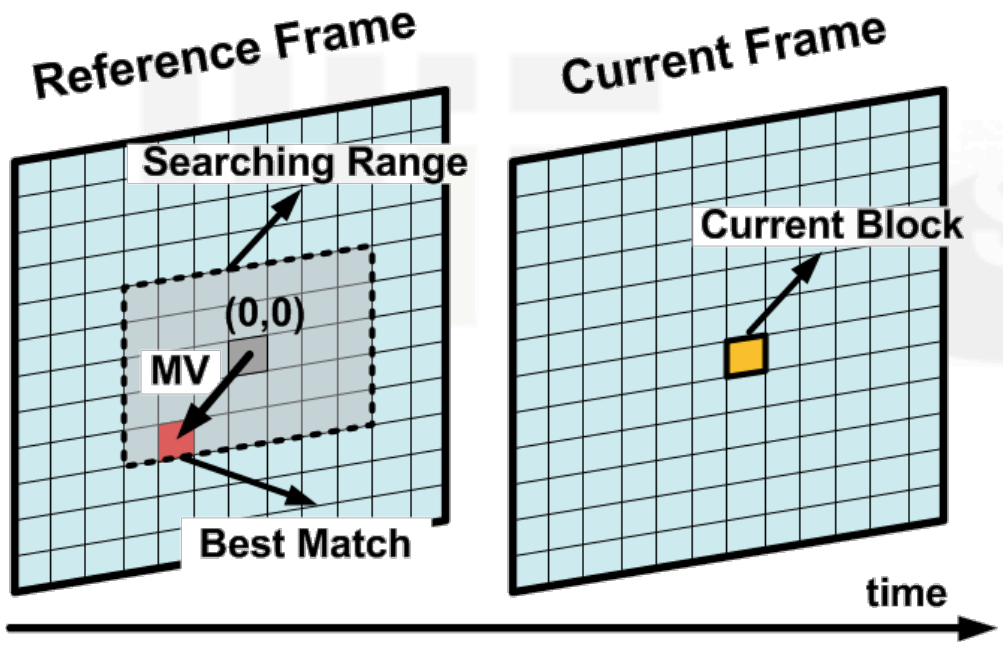


- For Intra-coded CU
  - PU options



Motion Estimation

- Search for block in reference frame(s) to predict current block with least rate-distortion cost
  - Signal block in previous frame using a motion vector
- Typically most computationally intensive function in encoder



Search algorithm considerations

1. Number of candidates
   - Number of computations
   - Number of memory accesses
2. Off-chip bandwidth
3. On-chip bandwidth

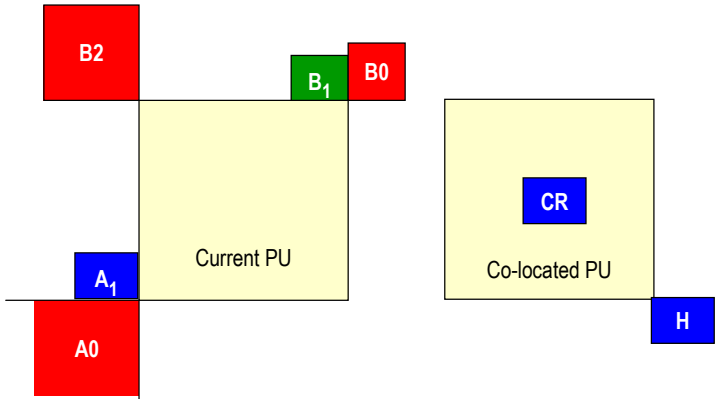
Motion Estimation in HM

- Integer pixel motion estimation
  - Rate is the bits required to transmit the motion data (including impact of motion predictor)
  - Distortion is calculated from the SAD of original and motion-compensated prediction (subsampled when block size > 8)

$$\underset{MV,\,REF}{\operatorname{argmin}} \sum_{i,j} \left| Diff(i,j) \right| + \lambda \cdot R(MV,\, REF)$$

where

- MV = motion vector (include impact of advanced mv predictor)
- REF = reference index



K. McCann et al "High Efficiency Video Coding (HEVC) Test Model 14 (HM 14) Encoder Description," JCTVC-P1002, 2014

Motion Estimation in HM

- Integer pixel motion estimation
  - Search Strategy
  1. Search center is motion vector predictor
  2. Diamond search around center (search range = 64 → 7 steps [1, 2, 4.. 64]); early termination if best candidate doesn't change in 3 steps.
  3. If best candidate > 5 pixels away from search center, do raster scan search (5 pixel steps).
  4. Perform diamond search around best candidate from step 2 or 3. If new best candidate found repeat 4.

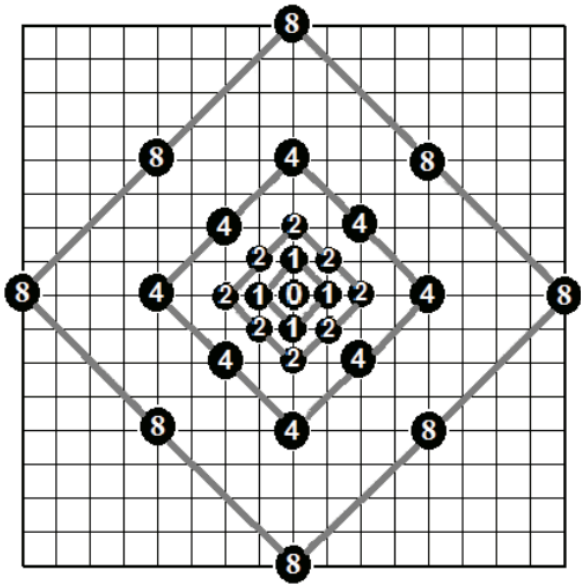


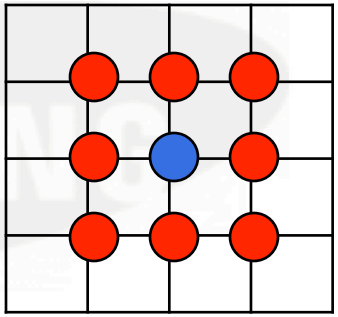
Image Source: N. Purnachand et al., IEEE ICCE-Berlin, 2012

Reference

- K. McCann et al "High Efficiency Video Coding (HEVC) Test Model 14 (HM 14) Encoder Description," JCTVC-P1002, 2014
- M. Sinangil, PhD Thesis, MIT, 2012

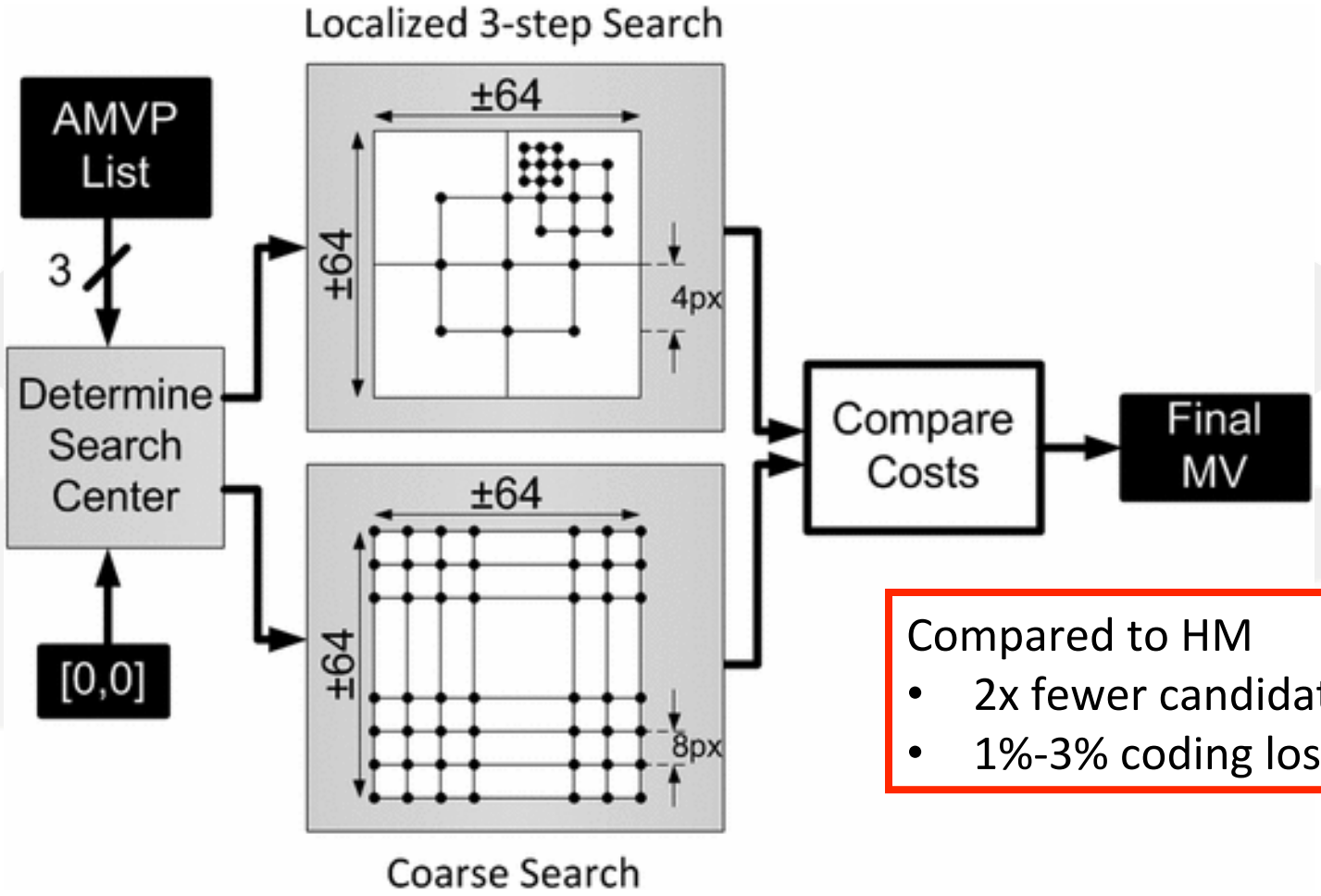
Motion Estimation in HM

- Half pixel motion estimation
  - Rate is the bits required to transmit the motion data (including impact of motion predictor)
  - Distortion is calculated from SATD
    - Block-wise 4x4 or 8x8 Hadamard transform on difference between original and motion-compensated prediction, and sum absolute coefficients
  - Search 8 points surrounding best integer motion vector
- Quarter pixel motion estimation
  - Same rate and distortion calculation as half pixel
  - Search 8 points surrounding **best half** pixel motion vector
- Also do search for merge/skip candidates



K. McCann et al "High Efficiency Video Coding (HEVC) Test Model 14 (HM 14) Encoder Description," JCTVC-P1002, 2014

Multiple Searches in Parallel

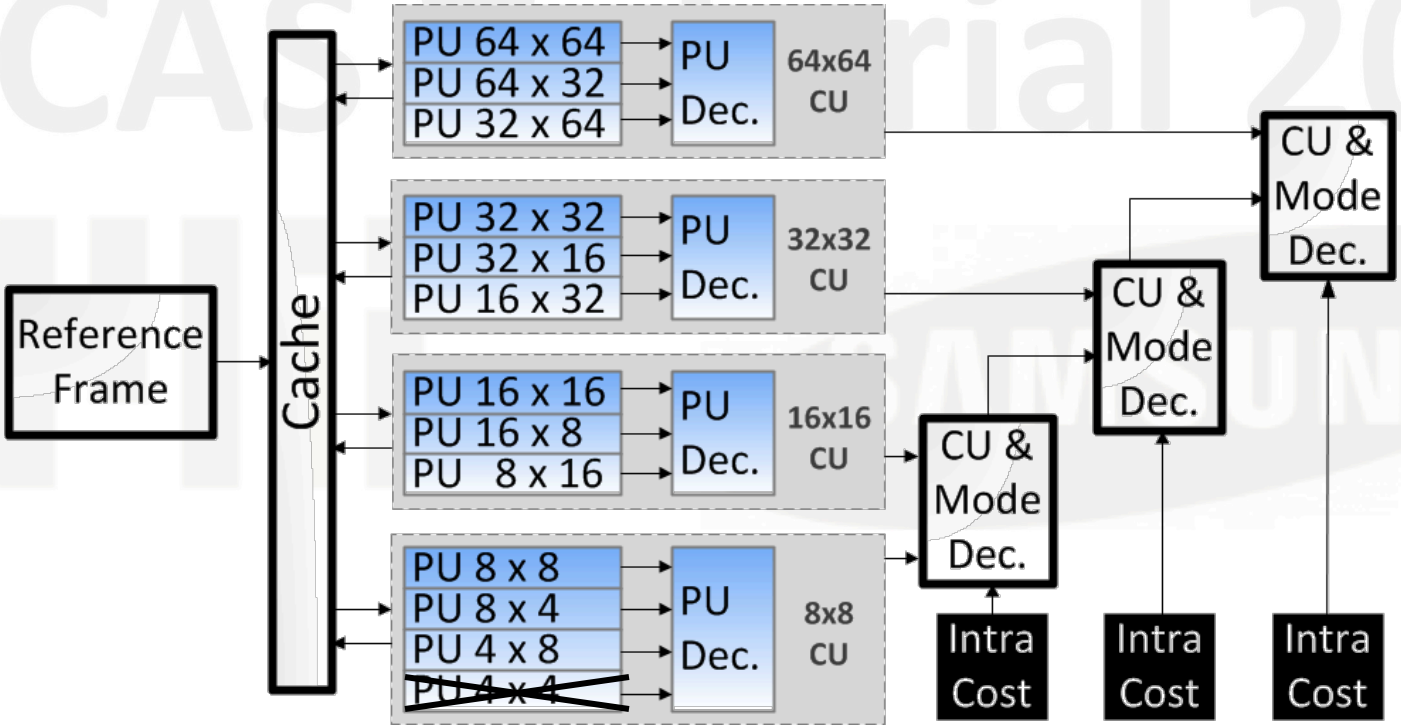


Compared to HM
- 2x fewer candidates
- 1%-3% coding loss

M. E. Sinangil et al., "Cost and Coding Efficient Motion Estimation Design Considerations for High Efficiency Video Coding (HEVC) Standard," *IEEE Journal of Selected Topics in Signal Processing*, 2013.

Parallel Motion Estimation

- Perform motion estimation for each PU in inter-coded CU
- Process CUs in parallel to increase throughput
  - Share search pixels across engines to reduce memory bandwidth by 8x



M. E. Sinangil et al., "Cost and Coding Efficient Motion Estimation Design Considerations for High Efficiency Video Coding (HEVC) Standard," *IEEE Journal of Selected Topics in Signal Processing*, 2013.

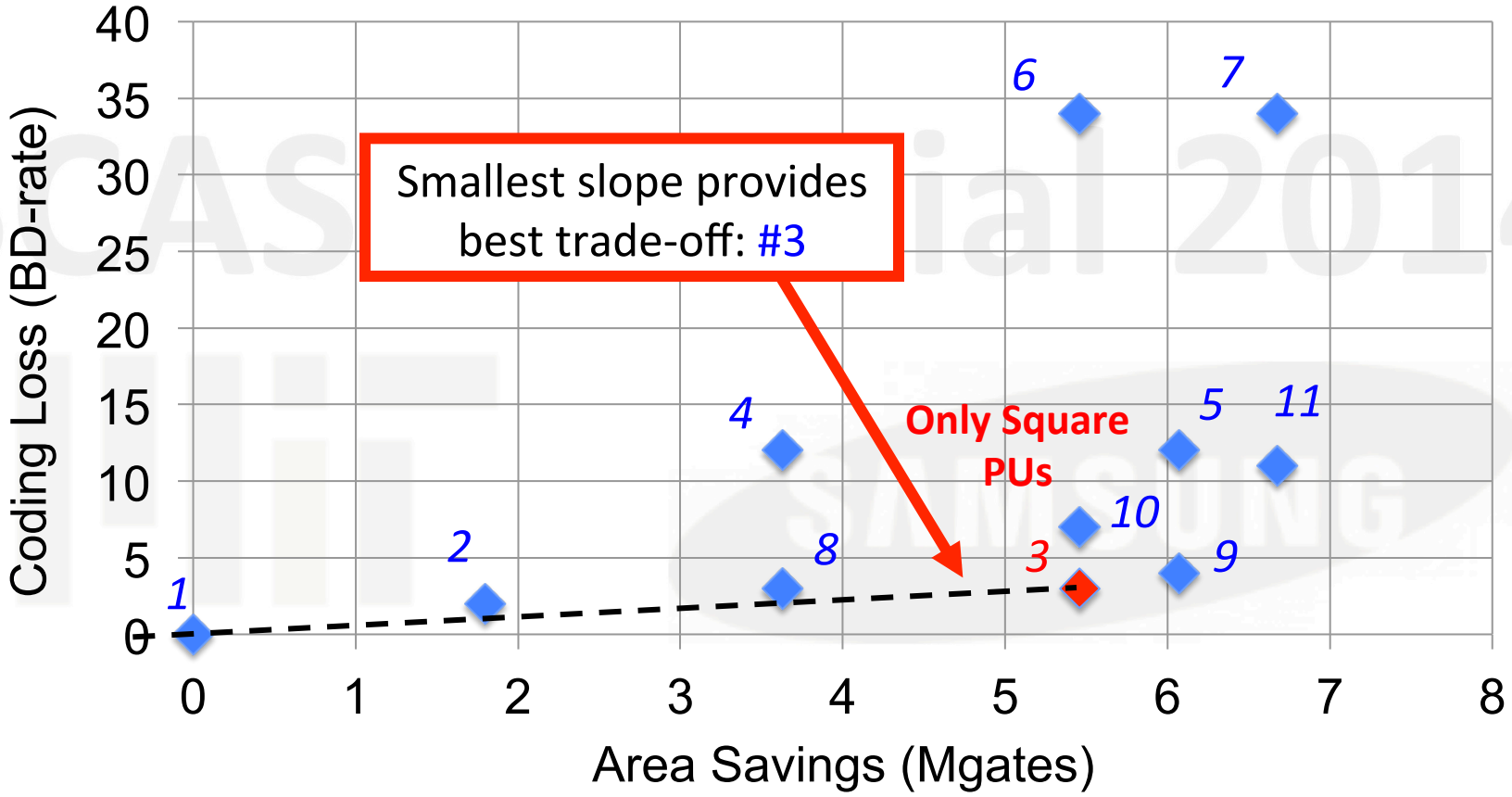
Reduce Number of PUs Processed

	Configuration #										
	1	2	3	4	5	6	7	8	9	10	11
64x64	Y	Y	Y	Y	Y	Y	Y	N	N	N	N
64x32	Y	Y	N	Y	N	Y	N	N	N	N	N
32x64	Y	Y	N	Y	N	Y	N	N	N	N	N
32x32	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N
32x16	Y	Y	N	Y	N	N	N	Y	N	N	N
16x32	Y	Y	N	Y	N	N	N	Y	N	N	N
16x16	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y
16x8	Y	Y	N	N	N	N	N	Y	N	Y	N
8x16	Y	Y	N	N	N	N	N	Y	N	Y	N
8x8	Y	Y	Y	N	N	N	N	Y	Y	Y	Y
8x4	Y	N	N	N	N	N	N	N	N	N	N
4x8	Y	N	N	N	N	N	N	N	N	N	N
4x4	Y	N	N	N	N	N	N	N	N	N	N
Ref. Buffer Size (KB)	680	565	248	439	208	234	163	356	170	201	115
On-Chip BW (GB/s)	1581	429	209	121	59	32.5	17.3	409	205	351	192
Off-Chip BW (GB/s)	159	69	30.2	27.4	12.7	8.5	5.1	64	28.7	49.1	25.1
Bit-Rate Increase (%)	0	2	3	12	12	34	34	3	4	7	11

M. E. Sinangil et al., "Cost and Coding Efficient Motion Estimation Design Considerations for High Efficiency Video Coding (HEVC) Standard," *IEEE Journal of Selected Topics in Signal Processing*, 2013.

Number of Partition Units

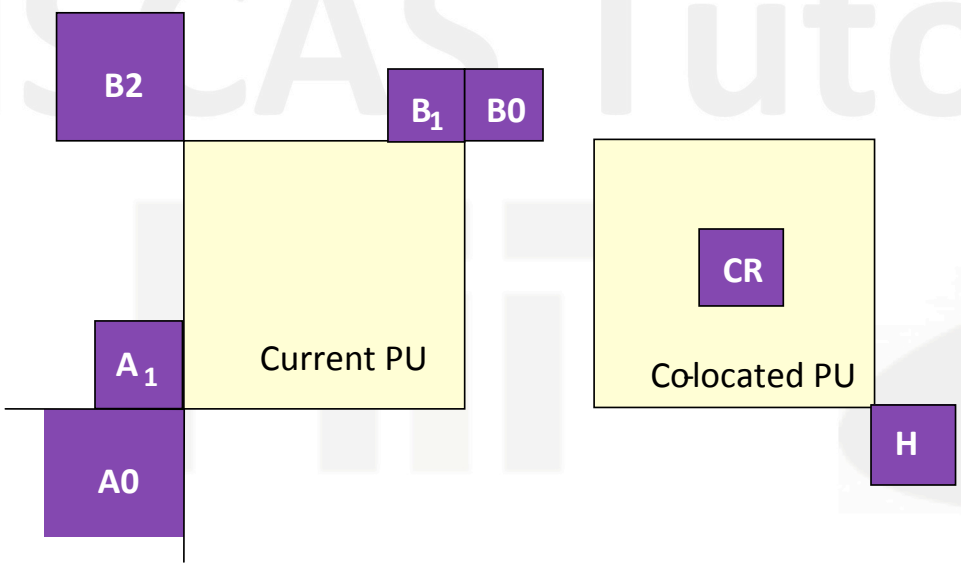
Trade-off between coding efficiency (BD-rate) and complexity (area cost) for different number of inter predicted partitions units

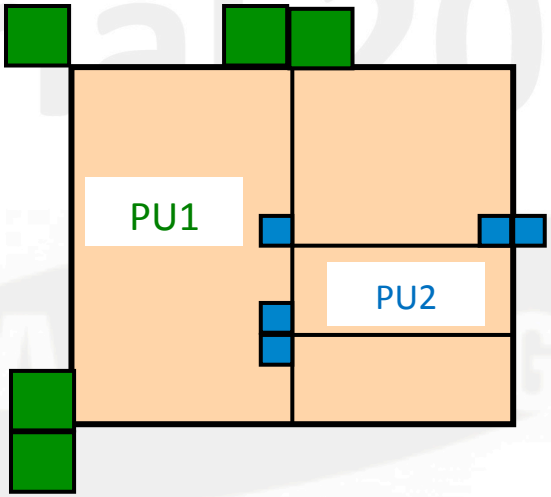


M. E. Sinangil et al., "Cost and Coding Efficient Motion Estimation Design Considerations for High Efficiency Video Coding (HEVC) Standard," *IEEE Journal of Selected Topics in Signal Processing*, 2013.

Motion Estimation with CU

- In HM, motion estimation done serially for PU within CU to get AMVP for accurate rate estimate

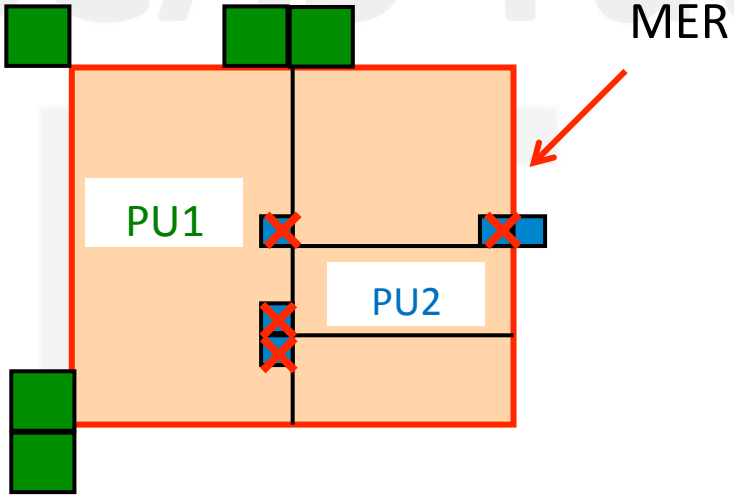




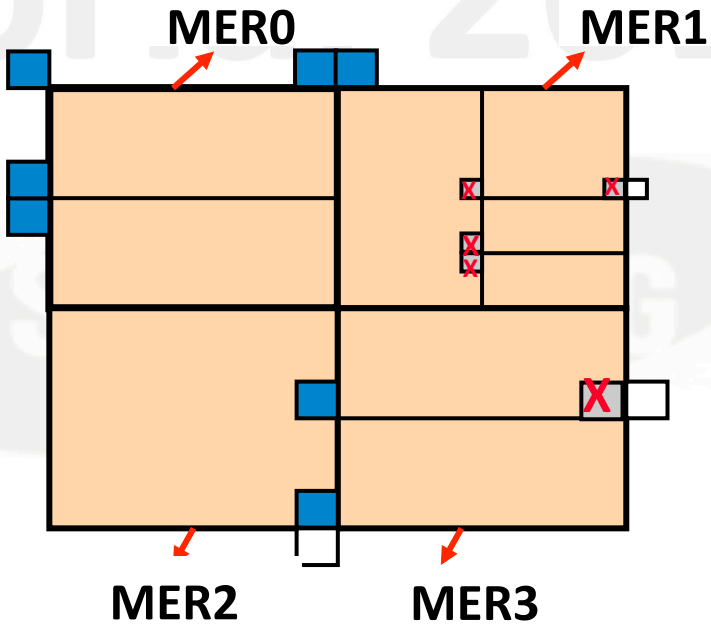
Can't process PU1 and PU2 in parallel

Parallel Motion Estimation

- HEVC has "Parallel Motion Estimation" feature to turn off dependency within an Motion Estimation Region (MER)
  - PU within region cannot use data from other PU in region
  - All PUs in region can be processed in parallel at encoder



Can process PU1 and PU2 in parallel



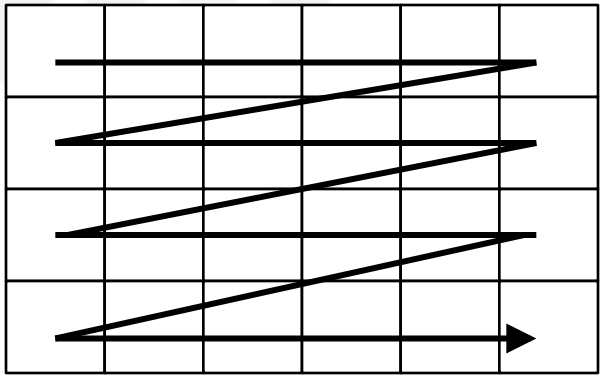
Multiple MERs per CTU

M. Zhou, "Parallelized merge/skip mode for HEVC," JCTVC-F069, 2011

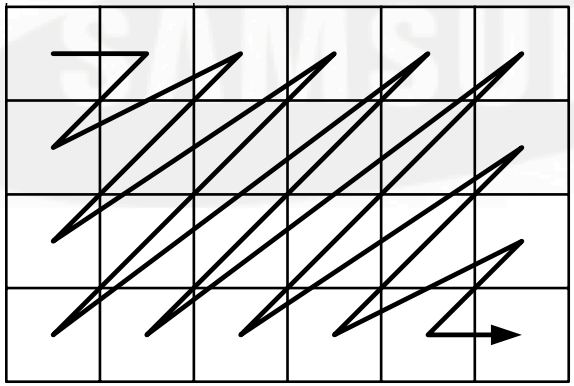
CTU Processing Order

- In HM, CTU processed in raster scan order
- Change CTU Processing Order to reduce reads from picture buffer (off-chip memory bandwidth) due to increased data locality
- Requires frame decoupling with entropy encoder (as entropy encoder must generate bitstream in raster scan order to be standard compliant)

Raster Scan



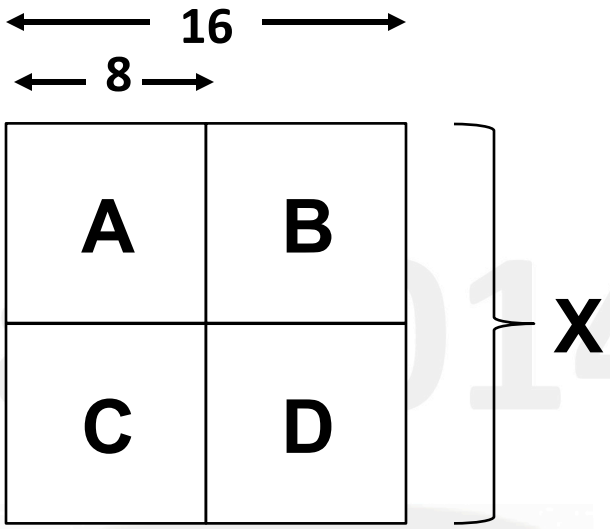
Alternative Scan



S. -F. Tsai et al., "Encoder Hardware Architecture for HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014.

Additional Complexity Reductions

- Bottoms up approach
  - Derive distortion cost for PU from sub-PUs (e.g. compute distortion of 16x16 PU from four 8x8 PU)
  - Requires storage of SAD sub-PUs
- Reduce bit-width for distortion calculation
- Use bilinear interpolation for fractional motion estimation

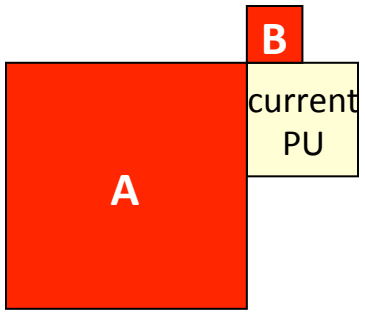


SAD16(X) =
SAD8(A) + SAD8(B) +
SAD8(C) + SAD8(D)

Intra Prediction Search in HM

- Rough mode decision: select N best mode out of 35
  - N equals 8 for 4x4, 8x8
  - N equals 4 for 16x16, 32x32, 64x64
  - Hadamard Cost Ranking (SATD distortion and mode bits for rate)
- Determine three Most Probable Modes (MPM)
  - Spatial neighbors to the left (A) and above (B)
  - If neighbors not available or redundant (A=B), use DC, Planar, vertical or adjacent angles (+/- 1)
- Decide between rough mode + MPM candidates
  - Full RDO (SSD for distortion and mode + coefficient bits for rate)

Y. Piao et al., "Encoder Improvement of Unified Intra Prediction," JCTVC-C207, Oct. 2010.



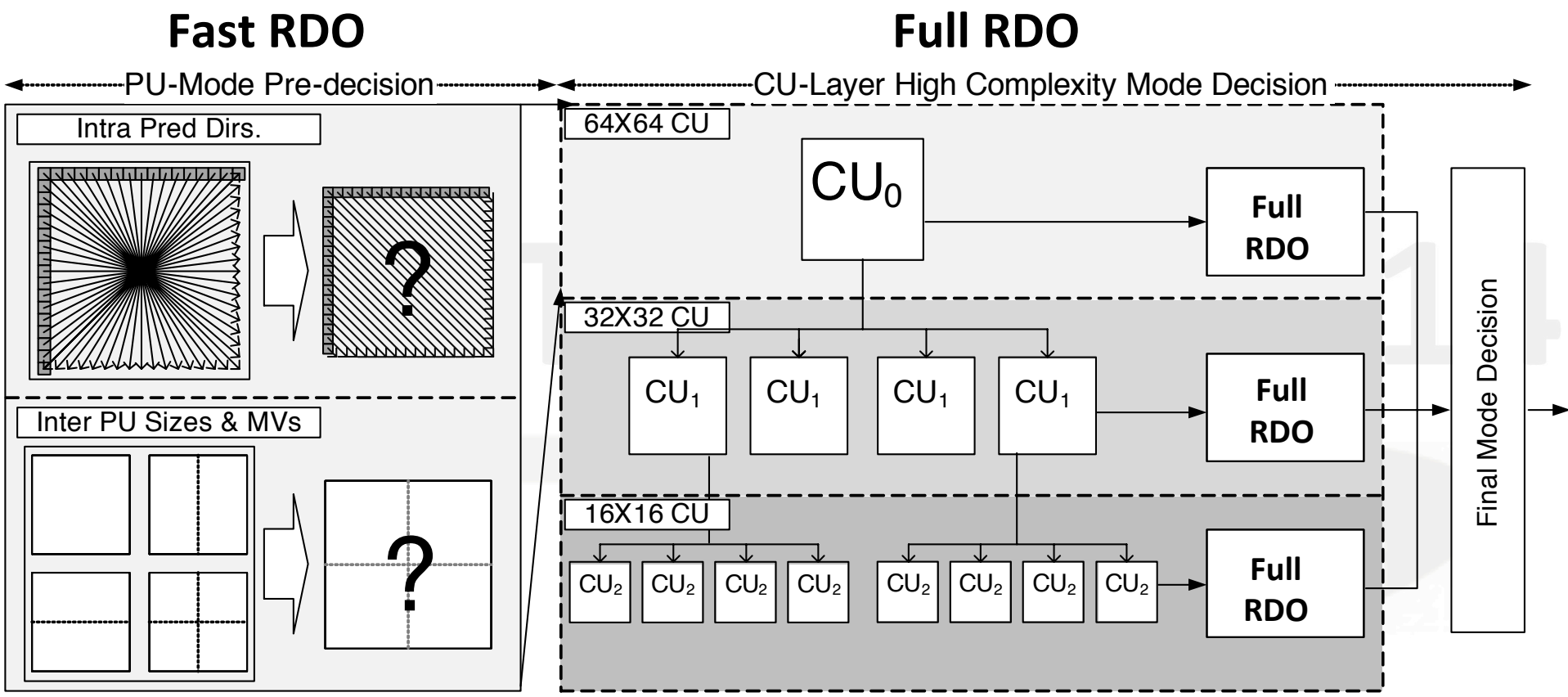
Additional Complexity Reduction

- To reduce search space, use coarse search with angular prediction, then refinement around coarse angles
- Skip 64x64 PU size
  - Since max TU is 32x32, prediction done at 32x32; thus only benefit of 64x64 intra-PU is signaling
- To increase throughput, use original pixels for intra prediction (rather than reconstructed pixels) to avoid dependence on reconstruction feedback loop

Above techniques have cumulative coding loss of 1%

S. -F. Tsai et al., "Encoder Hardware Architecture for HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014.

Hardware-Friendly RDO Pipeline

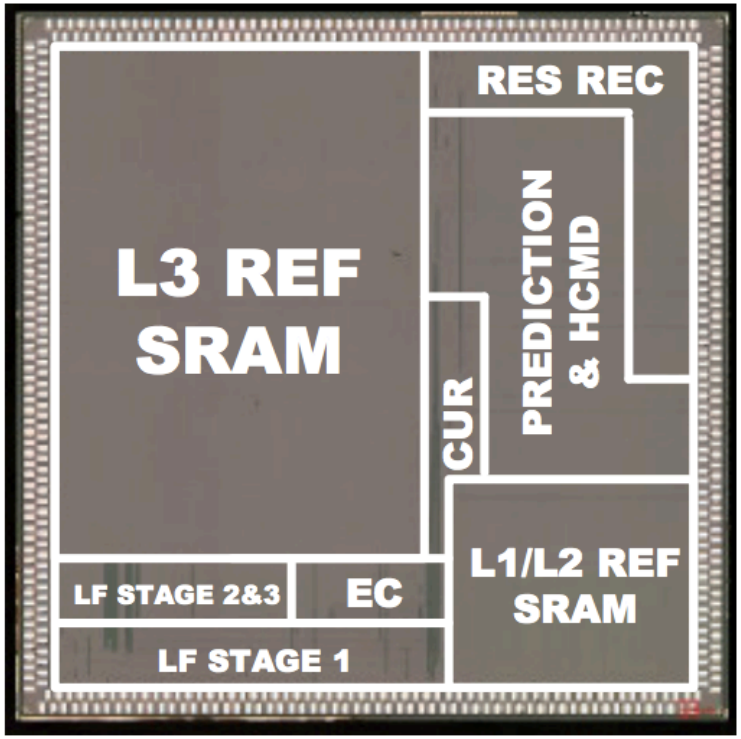


Only do full RDO on best Inter and Intra mode for each CU-depth (6% coding loss)

S. -F. Tsai et al., "Encoder Hardware Architecture for HEVC," *High Efficiency Video Coding (HEVC): Algorithms and Architectures,* Springer, 2014.

Hardware HEVC Encoder

Video Coding Standard	HEVC (WD4)
Technology	TSMC 28-nm HPM
Core Area	5x5mm^2
Gate Count	8350k
On-Chip Memory (SRAM)	7.14 MB
Resolution / Frame Rate	8192x4320@ 30fps
Frequency	312 MHz
Power	708 mW



S.-F. Tsai et al. , "A 1062Mpixels/s 8192×4320p High Efficiency Video Coding (H.265) encoder chip," IEEE VLSIC, 2013

ASIC Encoder Comparison

	ISSCC'09[22]	VLSIC'12[6]	This Work
Resolution	4096x2160@24fps	7680x4320@60fps	8192x4320@30fps
Throughput	212Mpixels/s	1991Mpixels/s	1062Mpixels/s
Standard	H.264 High @ Level 5.1	H.264 Intra	HEVC
Search Range	[-255,+255]/[-255,+255]	N/A	[-512,+511]/[-128,+127] (Predictor Centered)
Technology	TSMC 90nm	e-Shuttle 65nm	TSMC 28nm HPM
Core Size	3.95x2.90mm2	3.95x2.90mm2	5x5mm2
Gate Count	1732K	678.8K	8350K
Power	522mW@280MHz	139.9mW@280MHz	708mW@312MHz

S.-F. Tsai et al. , "A 1062Mpixels/s 8192×4320p High Efficiency Video Coding (H.265) encoder chip," *2013 Symposium on* VLSIC, 2013

Part IV: Emerging applications and HEVC extensions

What's Next

- More compression efficiency
  - Yes, in 5-10 years. Especially since video delivery is moving from traditional broadcast model to IP delivery and one-to-one streaming
  - Analogy: Public transport versus individual cars

Dallas
High Five

- Other considerations have become important too:
  - Power consumption, complexity, throughput
  - Ability to support new functionalities, modalities etc.

Changing Landscape of Video Coding Applications (1)

- Need for supporting diverse clients with varying capabilities (resolution, computational power etc.)

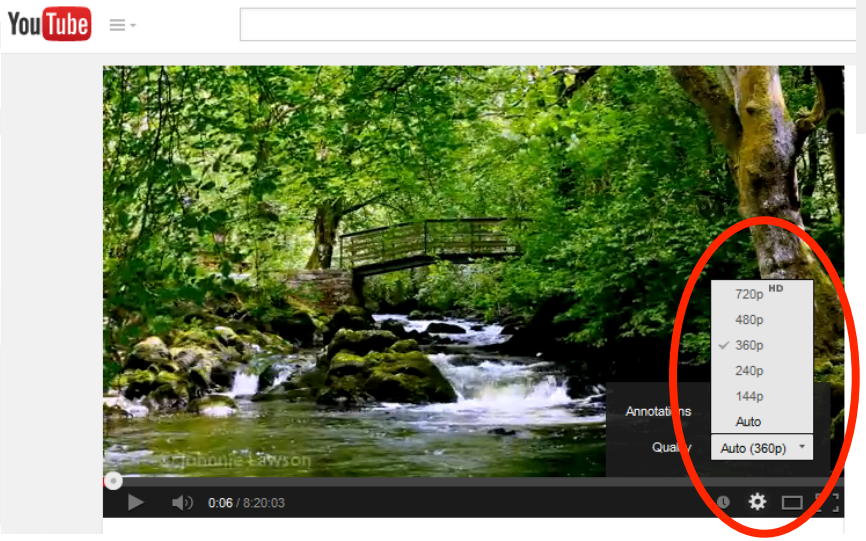


Image source: Samsung, Youtube

Changing Landscape of Video Coding Applications (2)

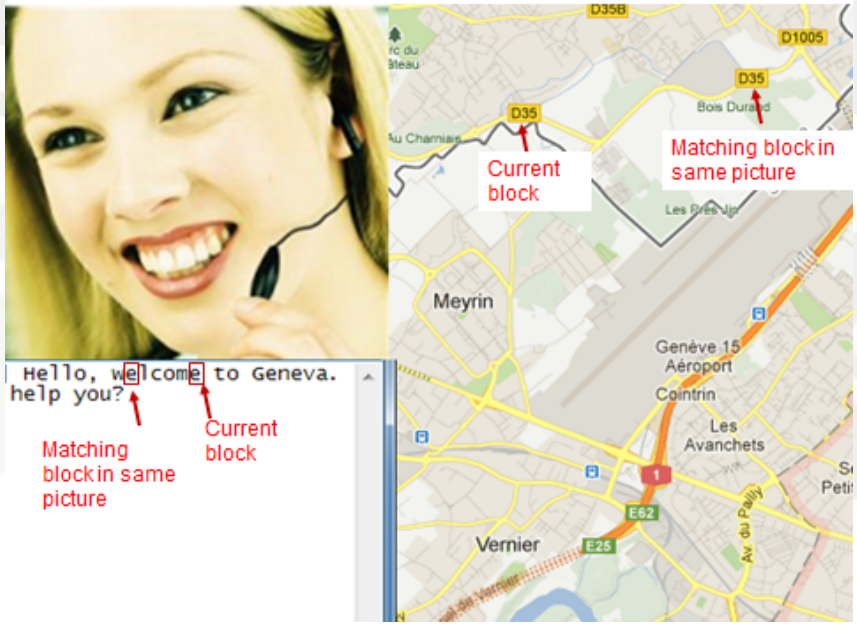
- Immersive experience
  - Multiple cameras and at higher video resolutions (1080p ➔ 4K ➔ 8K)
  - Multiple displays, Bigger displays (1080p ➔ 4K ➔ 8K)

  - Free-viewpoint video, 360degree video, augmented reality, 3D movies
  - Demos
    - http://replay-technologies.com/
    - http://www.kolor.com/video

Image source: Cisco, Kolor

Changing Landscape of Video Coding Applications (3)

- Growing requirement to support mixed format content consisting of natural video + graphics/text



Scalable Video Coding

Supporting Diverse Clients - Simulcasting

2560x19200

Encode

Bitstream 1

1280x960

Encode

Bitstream 2

Server

Client

640x480

Encode

Bitstream 3

Can we do better?

Scalable Video Coding

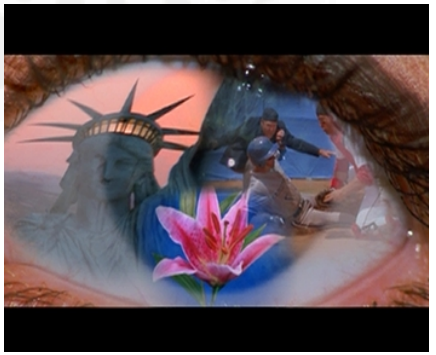
Spatial scalability

… 0110111 …

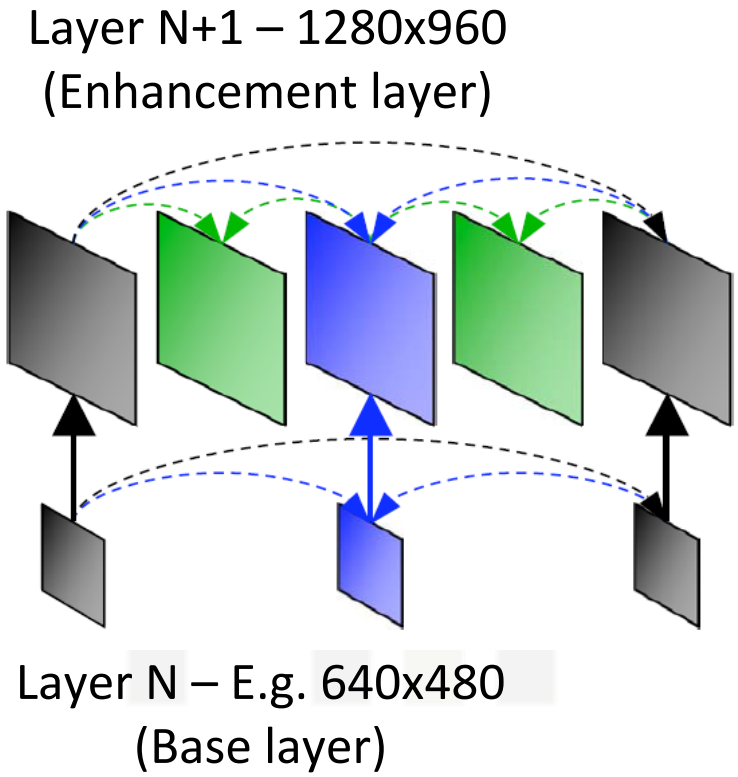
Single Bitstream

Temporal scalability

Quality (SNR) scalability



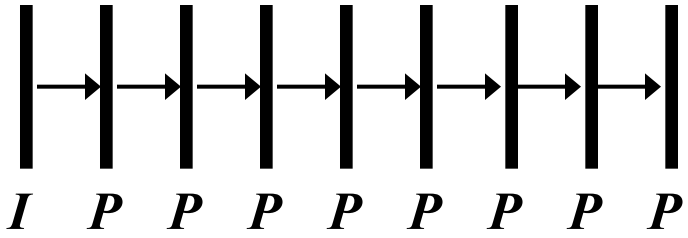
Spatial Scalability



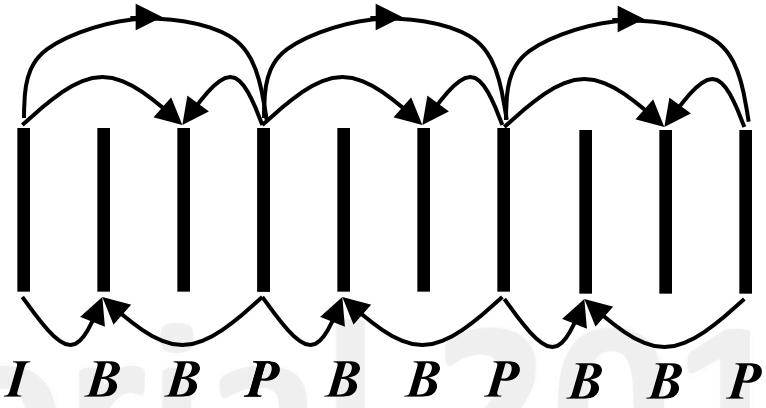
- Layered coding
- Higher layers have higher spatial resolution when compared to lower layers
- Upper layers re-uses data from lower layers

Figure source: T. Wiegand, JVT-W132 [1].

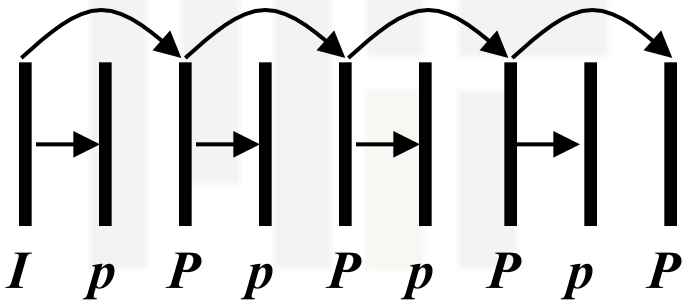
Temporal Scalability



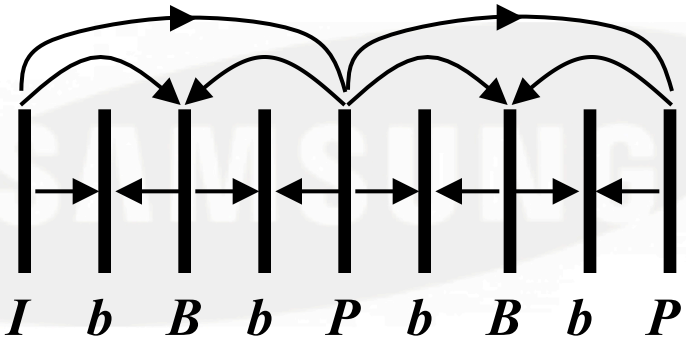
IPPP coding



IBBP coding



Hierarchical P-frames

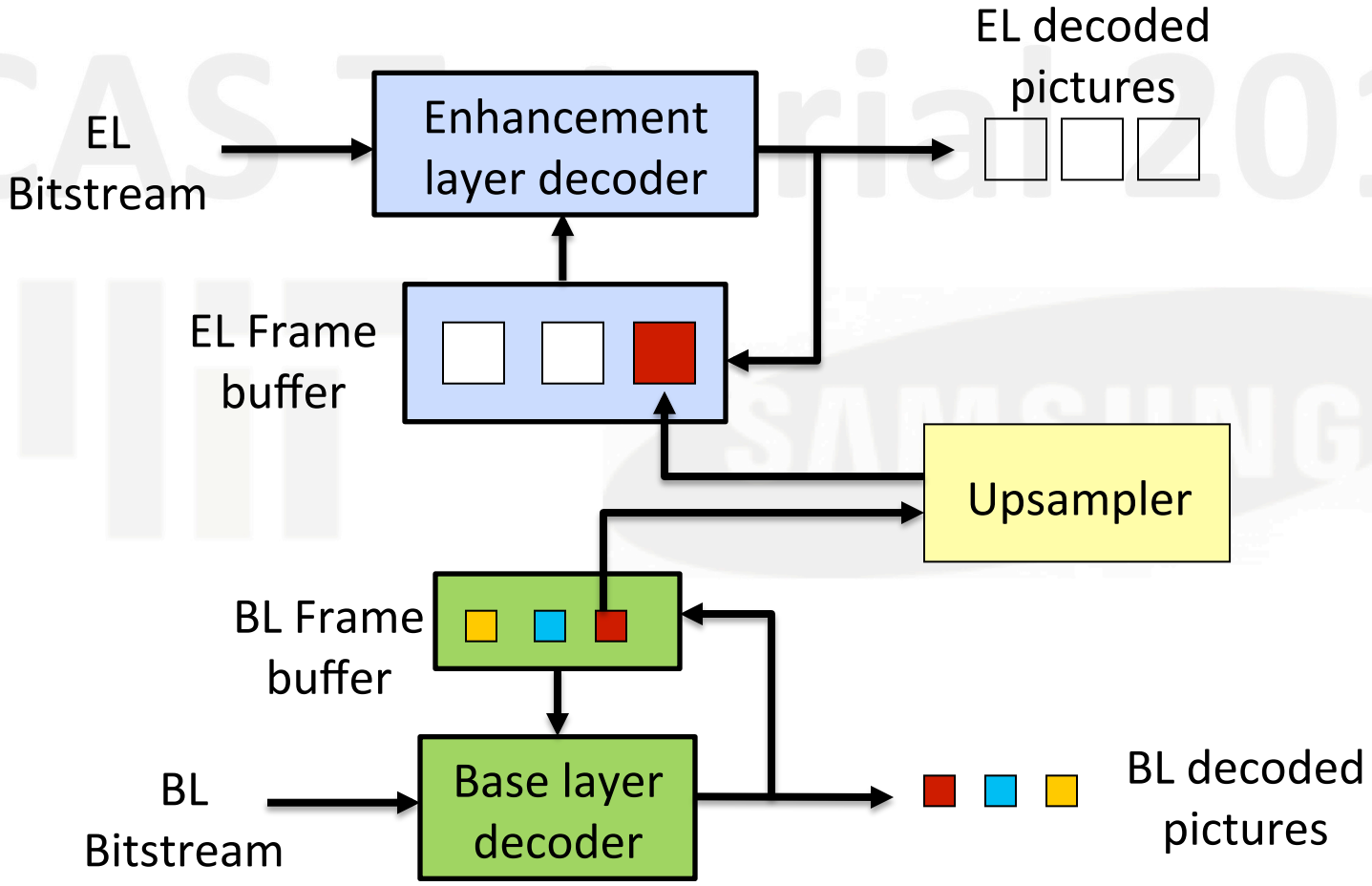


Hierarchical B-frames

- p, b – Non-reference frames

HEVC Scalable Extension (SHVC)

- SHVC: Scalable extension: Expected July 2014
- EL – Enhancement layer, BL – Base layer



SHVC Performance

- 2x scalability (i.e. base layer is half the size of enhancement layer) compared to simulcast

Coding configuration	BD-Rate savings
All Intra coding	23%
Random access (Hierarchical-B)	16%

- Quality (SNR) scalability compared to simulcast

Coding configuration	BD-Rate savings
All Intra coding	28%
Random access (Hierarchical-B)	20%

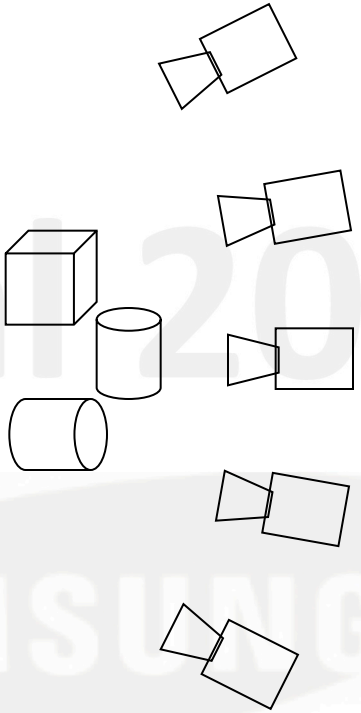
D.-K. Kwon, M. Budagavi, "Combined scalable and mutiview extension of High Efficiency Video Coding (HEVC)," *IEEE Picture Coding Symposium,* pp. 414 – 417, 2013.

Multiview Video Coding

Multiview Video Capture

Stereo, 3D video

360degree video



Free viewpoint video

Image source: Fuji, Kolor

Stereoscopic Video Coding

Left View

Right View

Stereo Video encoding

Stereo video bitstream

Camera modules

Stereo video bitstream

Stereo Video decoding

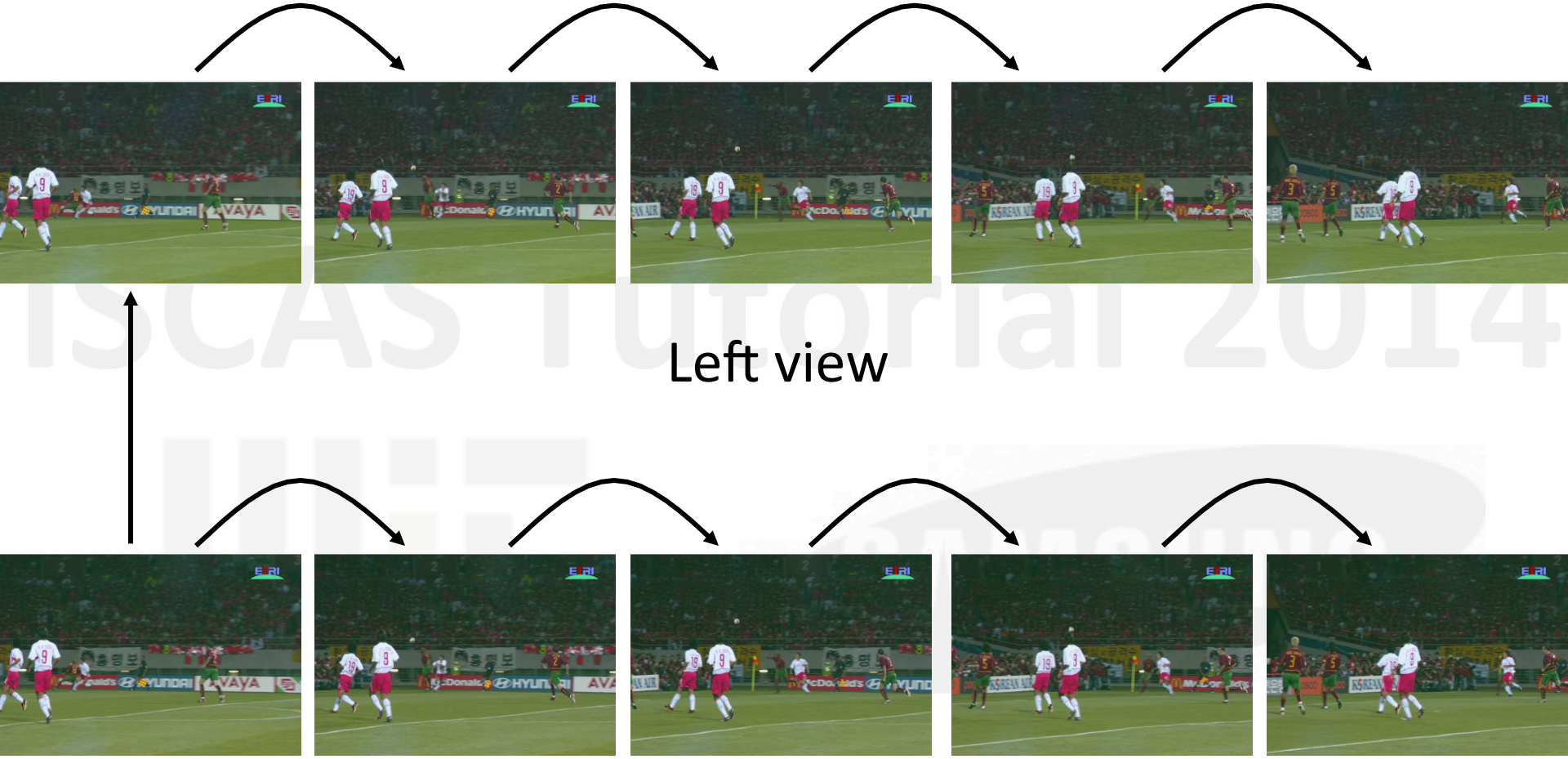
Left View

Right View

3D display

Image source: Samsung

Redundancy in Stereo Video

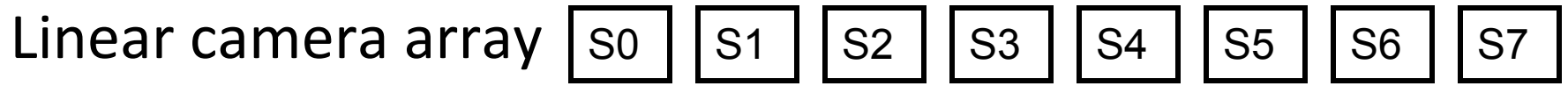


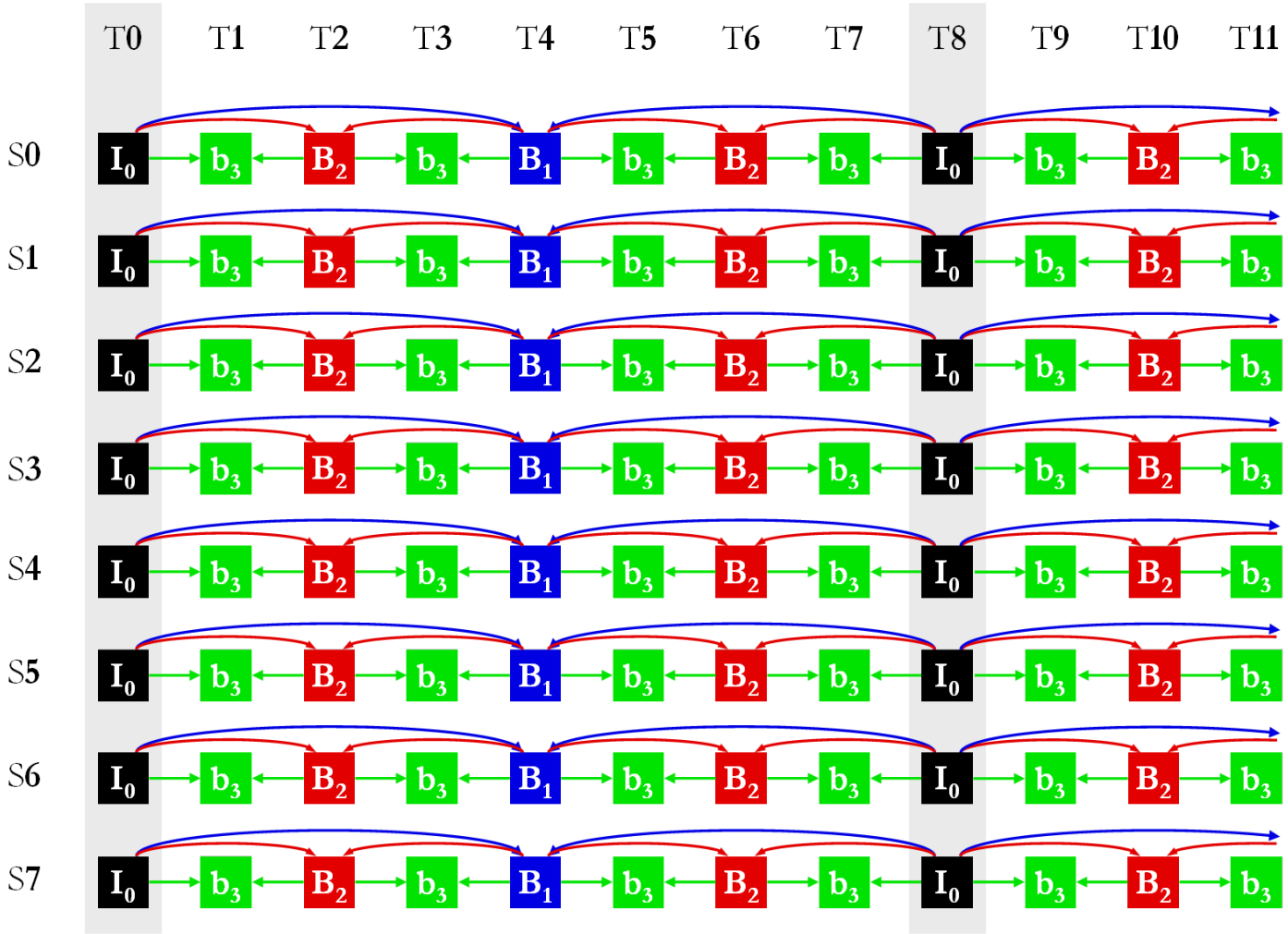
Left view

Right view

Multiview Video Coding – Picture Prediction Structures (1)

- Linear camera array

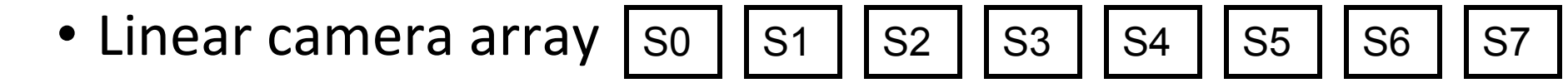


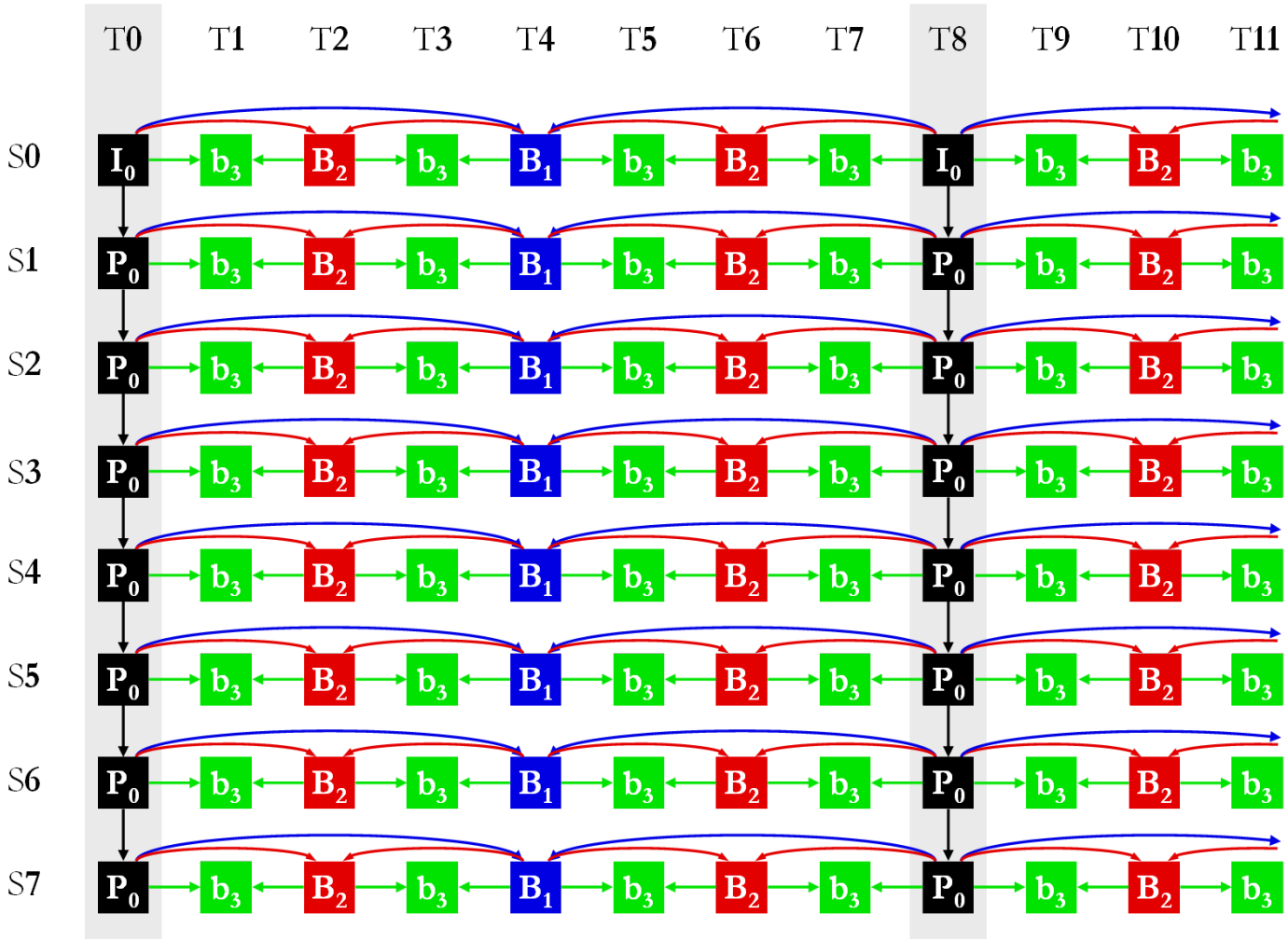


Simulcast

Multiview Video Coding – Picture Prediction Structures (1)

- Linear camera array

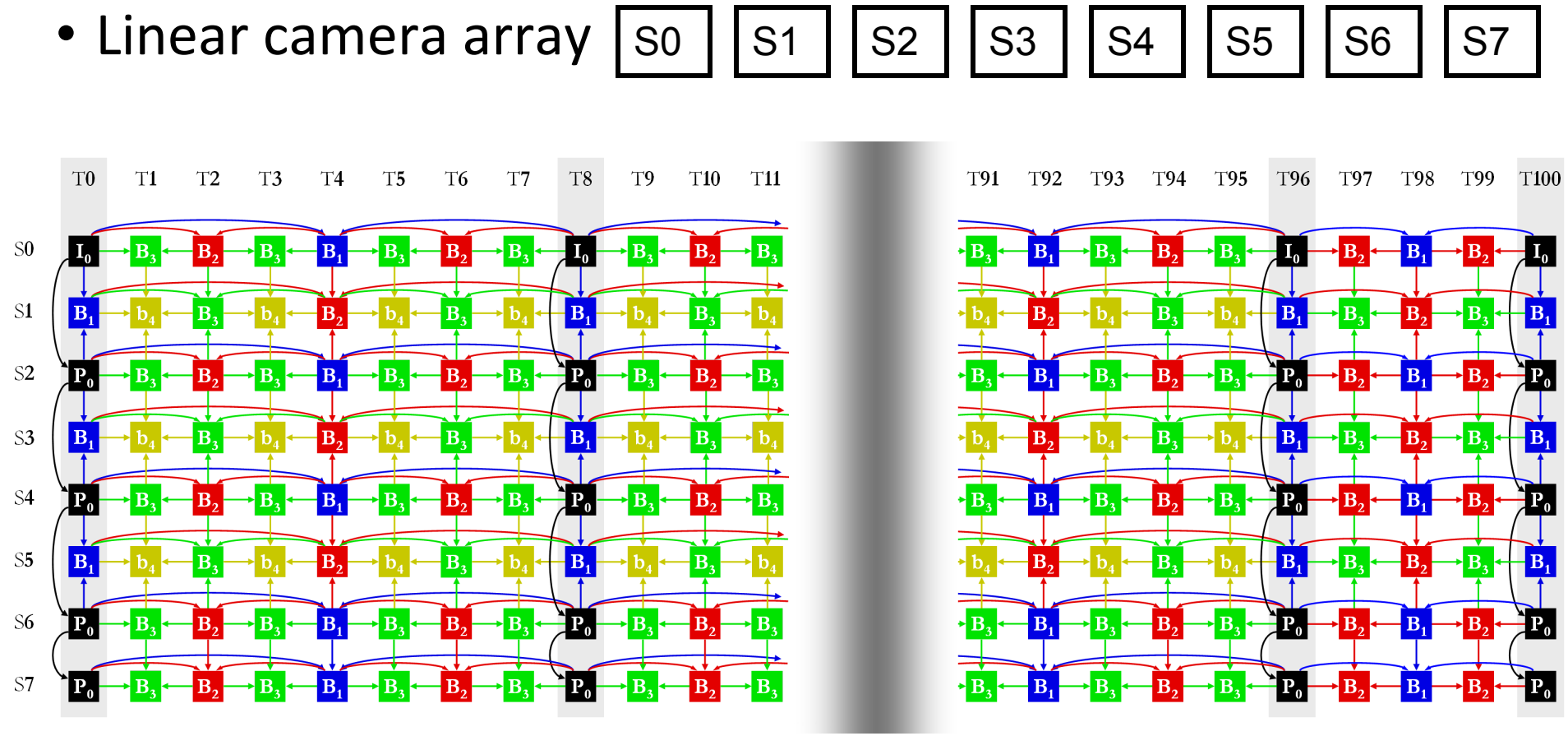




Interview prediction of anchor frames

Multiview Video Coding – Picture Prediction Structures (1)

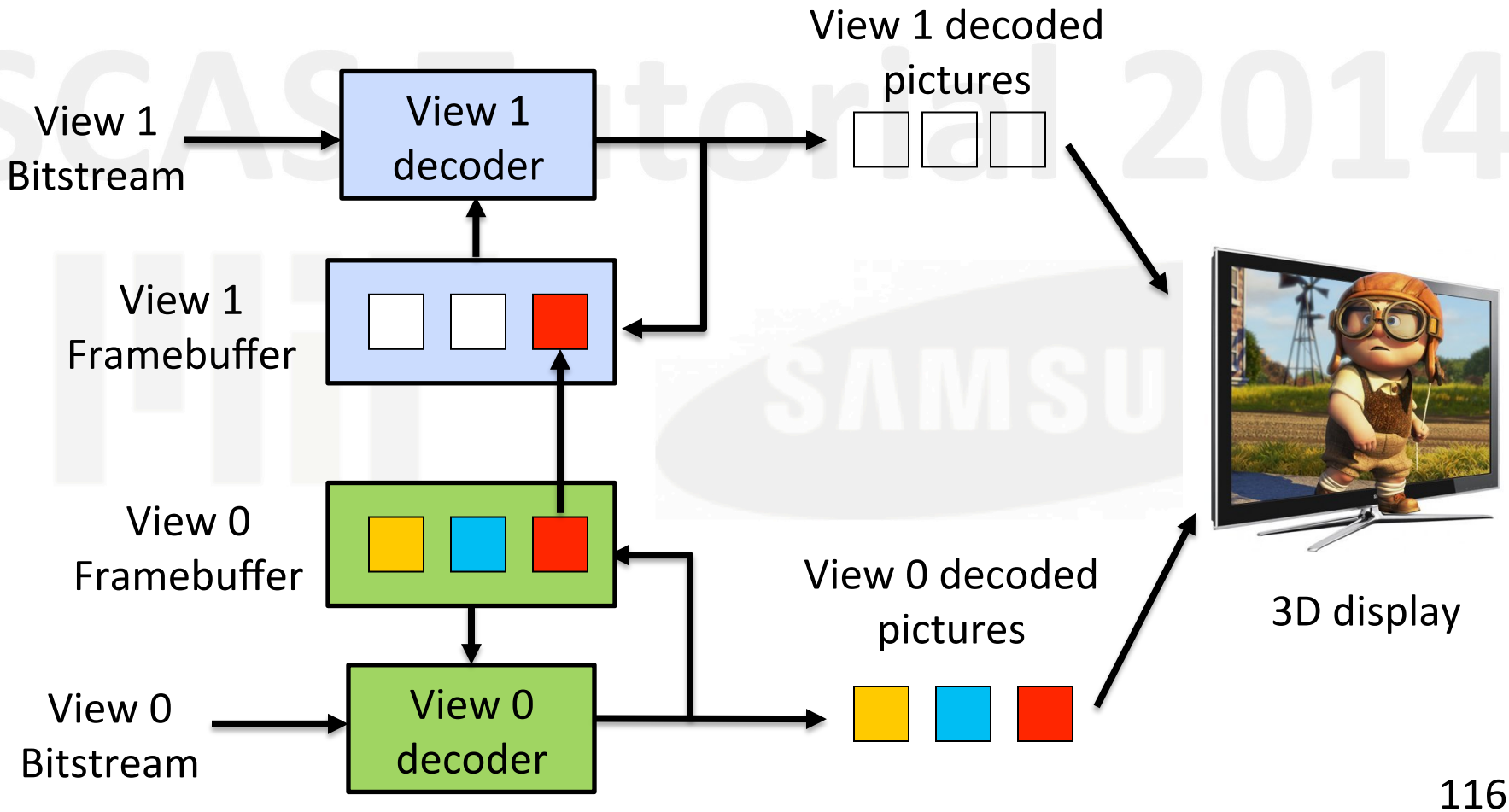
- Linear camera array



Both anchor and non-anchor views predicted from other views

HEVC Multiview Extension (MV-HEVC)

- MV-HEVC : Multiview extension: Expected July 2014
- View 0: Left view, View 1: Right view



Combined Scalable and Mutiview Extension of HEVC

- Applications of the combined scalable and multiview HEVC coding include:
  - Scalable stereoscopic video (e.g. 1080p stereo to the emerging 4K stereo),
  - Mixed resolution multiview coding
- H.264/AVC does not support combined scalable and multiview coding
- HEVC allows for combined scalable and multiview coding

D.-K. Kwon, M. Budagavi, "Combined Scalable and Mutiview Extension of High Efficiency Video Coding (HEVC)", *IEEE Picture Coding Symposium*, 2013.

Combined Scalable and Mutiview Extension of HEVC

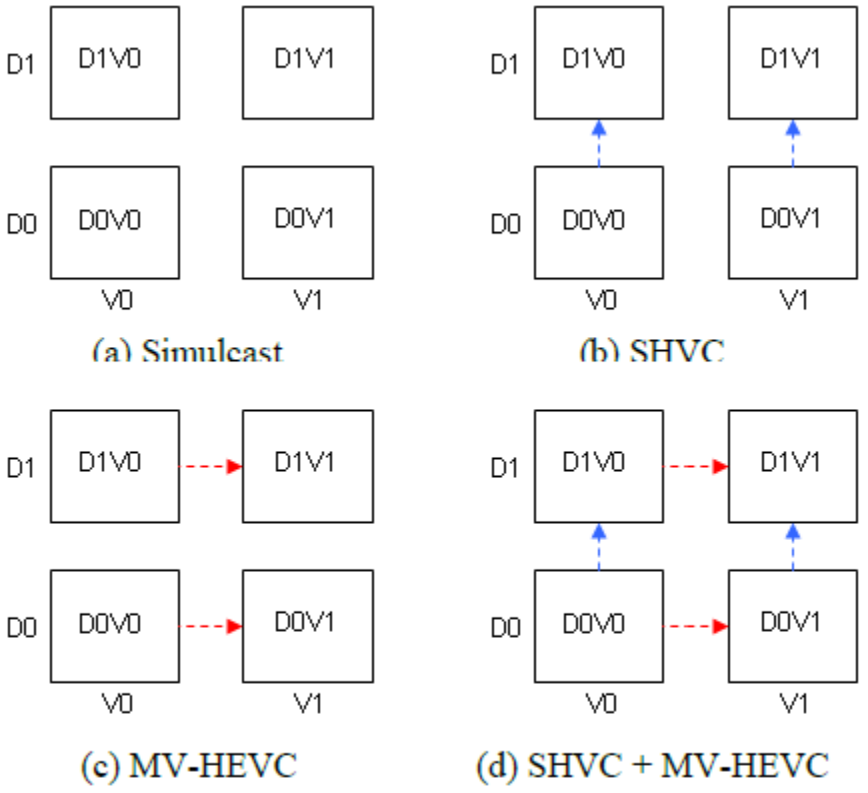


Figure 3. Prediction structures between layers for scalable stereo HEVC coding

D.-K. Kwon, M. Budagavi, "Combined Scalable and Mutiview Extension of High Efficiency Video Coding (HEVC)", *IEEE Picture Coding Symposium*, 2013.

Combined Scalable and Mutiview Extension of HEVC

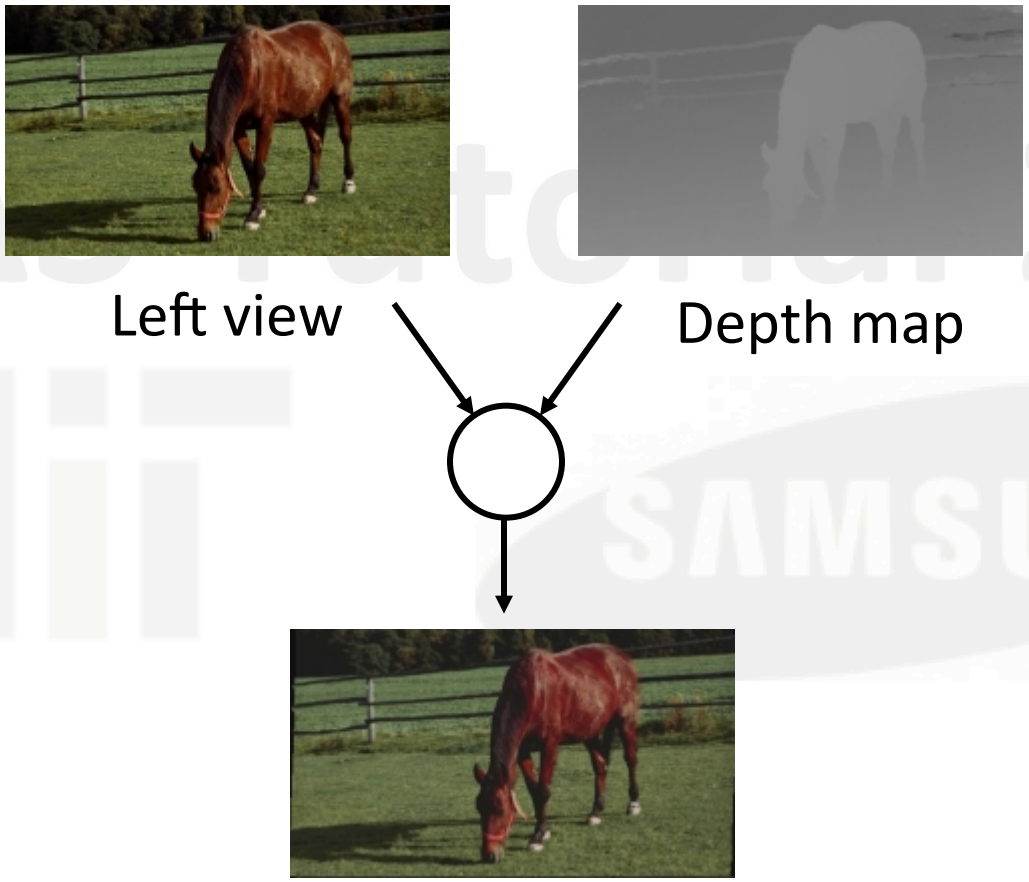
TABLE IV. 'BL-D + EL-D' BD-RATE (%) OF REFIDX SHVC + MV-HEVC W.R.T MV-HEVC.

	2x			SNR		
	Y	*Cb*	*Cr*	*Y*	*Cb*	*Cr*
AI	-19.5	-17.1	-17.5	-24.4	-22.3	-22.6
RA	-12.7	-5.0	-5.6	-16.4	-7.8	-8.9
LDP	-7.9	-0.1	-1.5	-9.0	-1.8	-3.0

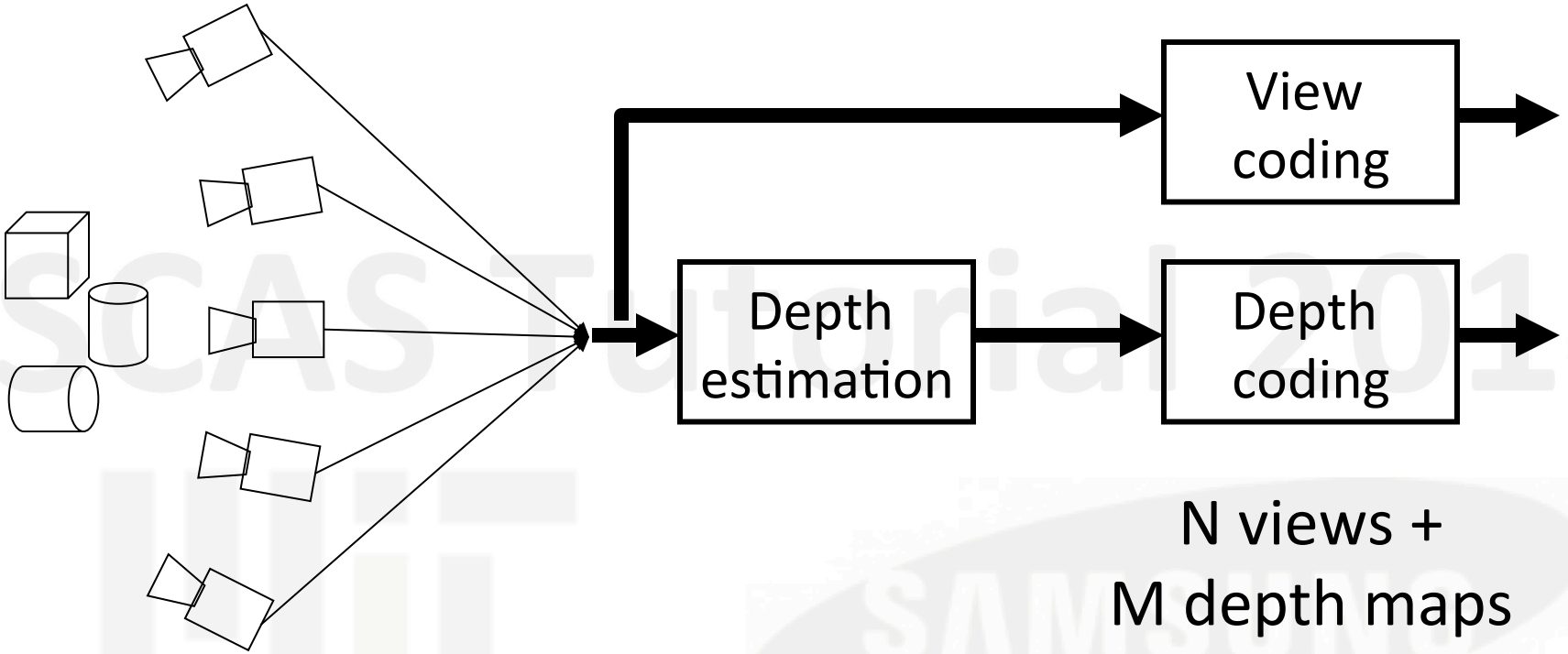
D.-K. Kwon, M. Budagavi, "Combined Scalable and Mutiview Extension of High Efficiency Video Coding (HEVC)", *IEEE Picture Coding Symposium*, 2013.

MV-HEVC + Depth (3D-HTM)

- Standardization in on-going

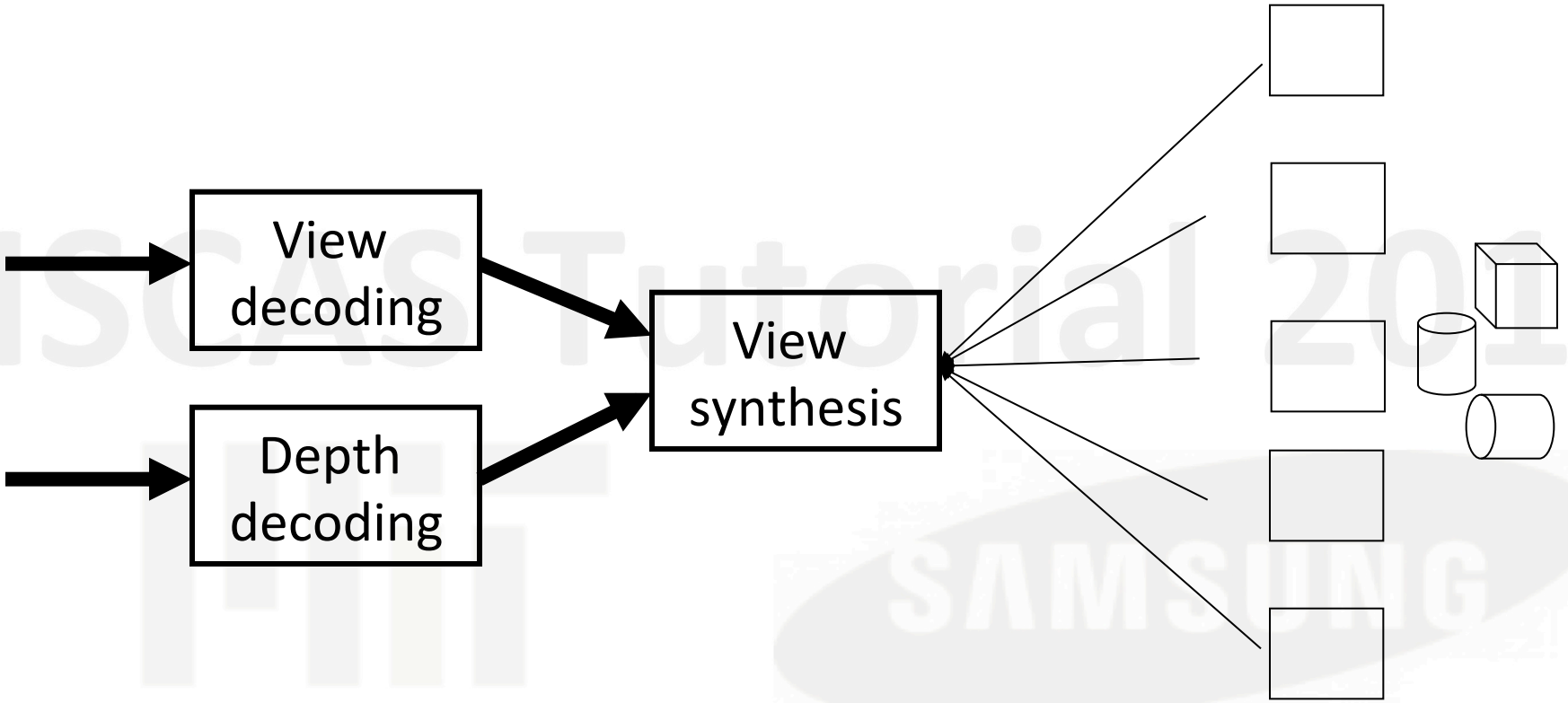


MV-HEVC + Depth Encoding



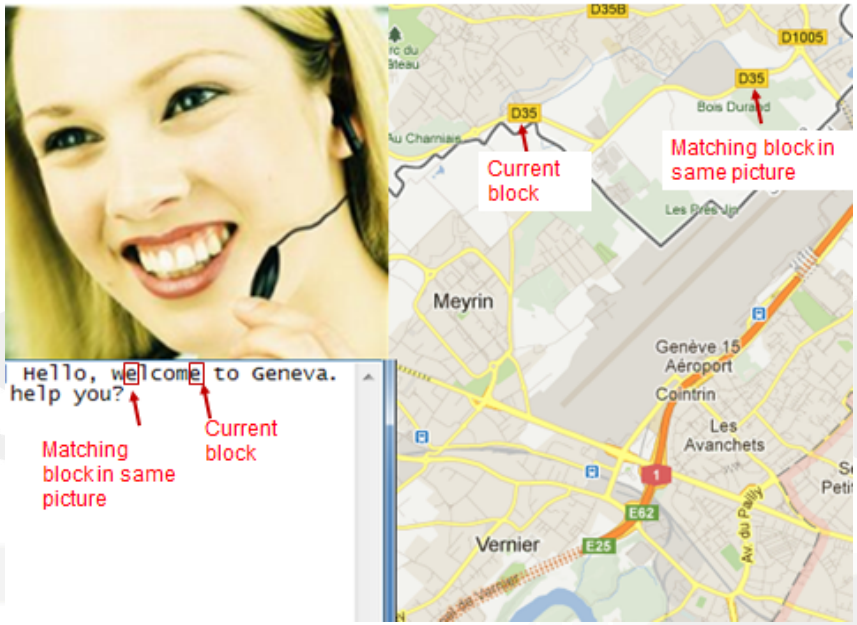
- Views that are transmitted will be coded using MV-HEVC
- Expect additional 20% gain

MV-HEVC + Depth Decoding



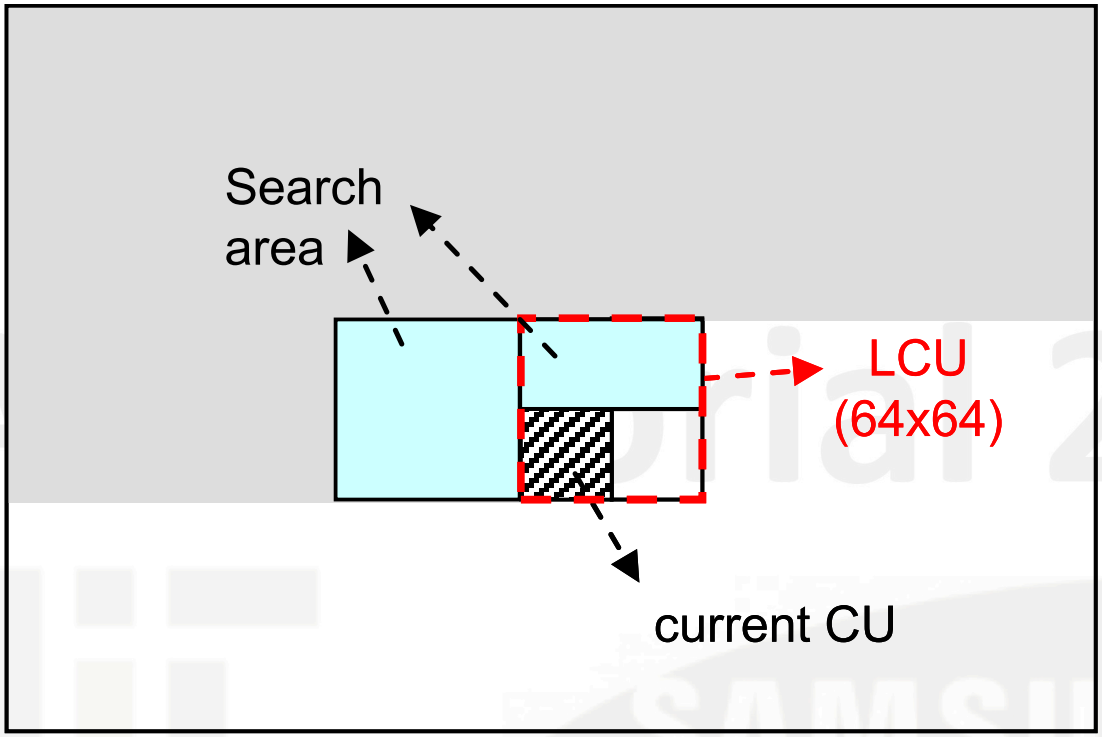
Screen Content Video Coding

Screen Content Coding



- Applications such as automotive infotainment, wireless displays, remote desktop, remote gaming, cloud computing etc. are becoming popular
- Video in these applications often has mixed content consisting of natural video, text, graphics etc.
  - In text and graphics regions, patterns (e.g. text characters, icons, lines etc.) can repeat within a picture
  - Also blocks with limited set of colors are possible

Intra Block Copy



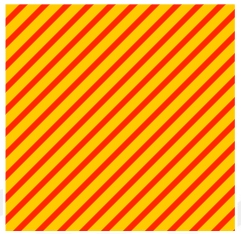
Bit-rate savings

	Intra	Random access	Low delay
SC RGB 444	27.0%	21.5%	17.0%
SC YUV 444	23.5%	20.2%	15.9%

M. Budagavi, D.-K. Kwon, "Intra motion compensation and entropy coding improvements for HEVC screen content coding", *IEEE Picture Coding Symposium*, 2013.

Palette Coding

- Input video:
  - 8 bits per pixel, per color component
  - 4x4 block: 8*3*16 = 384 bits



- Palette coding:
  - Color palette: 2 Colors in our example: 2*24 = 48 bits
  - Color index: 1 bit per pixel in our example: 16 bits
  - Total bits: 64 bits

Color 0			
Color 1			
i0	i1	i2	i3
i4	i5	i6	i7
i8	i9	i10	i11
i12	i13	i14	i15

- Note: This slide shows a very simple example for explaining purposes. Techniques being evaluated currently cab use more colors in palette and more bits for color index.

HEVC Screen Content coding

- HEVC Screen content coding activity
  - Started in April 2014
  - Expected completion early-mid 2015
- Key tools being studied
  - Intra Block Copy with extended search area
  - Palette based coding

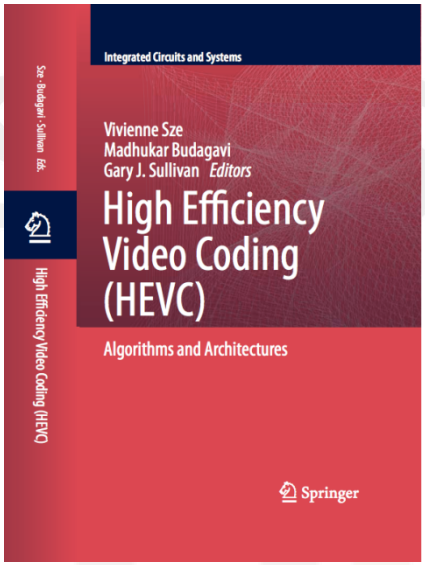
Summary

- Video content continues to impose a severe burden on today's global networks
  - Rapid growth in the usage and diversity of video applications and services
  - Increasing popularity of HD video and emergence of beyond-HD formats accompanied by stereo and multi-view content
- HEVC is the latest video coding standard, which gives 50% improvement in coding efficiency, and is expected to support video applications for the next decade.
- In addition to improving coding efficiency, implementation challenges were also considered to maximize processing speed and minimize hardware cost.

References

- V. Sze, M. Budagavi, G. J. Sullivan (Editors), "High Efficiency Video Coding (HEVC): Algorithms and Architectures," Springer, 2014
- G. J. Sullivan, et al. "Overview of the High Efficiency Video Coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology,* 2012
- J. Ohm et al., "Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC),"*IEEE Transactions on Circuits and Systems for Video Technology*, 2012

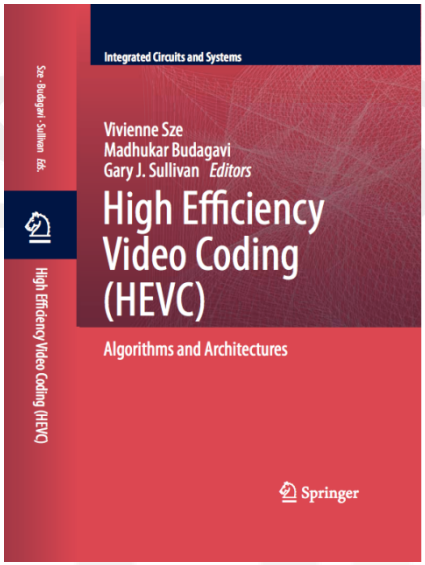
HEVC Book



- Introduction
- High-Level Syntax in HEVC
- Block Structures and Parallelism Features in HEVC
- Intra-Picture Prediction in HEVC
- Inter-Picture Prediction in HEVC
- Transform and Quantization in HEVC
- In-Loop Filters in HEVC
- Entropy Coding in HEVC
- Compression Performance Analysis in HEVC
- Decoder Hardware Architecture in HEVC
- Encoder Hardware Architecture in HEVC

http://www.springer.com/engineering/signals/book/978-3-319-06894-7

HEVC Book



The book serves the video engineering community by:

- Providing video application developers an invaluable reference to the latest video standard, High Efficiency Video Coding (HEVC);
- Serving as a companion reference that is complementary to the HEVC standards document produced by the JCT-VC – a joint team of ITU-T VCEG and ISO/IEC MPEG;
- Including in-depth discussion of algorithms and architectures for HEVC by some of the key video experts who have been directly involved in developing and deploying the standard;
- Giving insight into the reasoning behind the development of the HEVC feature set, which will aid in understanding the standard and how to use it.