WHAT EXACTLY IS THE DIFFERENCE BETWEEN MPEG-4 AVC AND MPEG-2?

MATTHEW GOLDMAN
VICE PRESIDENT TECHNOLOGY SOLUTION AREA TV, ERICSSON
ABSTRACT

› Many compression system applications have or will soon migrate from MPEG-2 to MPEG-4 AVC solutions. This session will explain the major differences between the two formats, from an overview of compression tools to transport considerations (MPEG-2 Transport Stream, IP, and file formats). Practical examples will be used where possible.

› Please note: TANDBERG Television is part of Ericsson now.
AGENDA

› Overview of MPEG-4 AVC

› Transport considerations

› Issues to consider
  – Concatenation of decode/encode stages
  – Legacy receiver devices
  – Ancillary data & audio
  – Contribution
MPEG-4 AVC VS. MPEG-2 VIDEO BIT-RATES

› HD: 40-60% more efficient than best MPEG-2
› SD: 30-50%
› Will continue to improve further

Examples*

- HD: MPEG-2 Video @13-18Mbps → MPEG-4 AVC @5-10Mbps
- SD: MPEG-2 @2.5-3.5Mbps → MPEG-4 AVC @1.2-2.5Mbps

*full resolution; varies greatly by content
OVERALL GOAL OF COMPRESSION

› Encode the smallest amount of information for a given “quality” level

› Remove spatial and temporal redundancies
  – Take advantage of deficiencies in the human visual/aural systems vs. pure mathematics

› Encode only residual or “error” signal
  – Frame “reconstruction” used to compare next frame with previous frame
  – Internal feedback loop in encoder matches decoder operation

› Better match between reconstructed frame and next frame = smaller residual = fewer bits to encode

› Encoder makes difficult decisions of how to compress efficiently (using the tool kit)
  – Decoder follows “directions” defined in received bit stream
PRINCIPLES OF COMPRESSION

› Encoders use the advantage of redundancy in moving images
  – Similarity within a frame (spatial redundancy)
  – Similarity between frames (temporal redundancy)

› Encoders create a prediction for the decoder to use
  – Then send the error signal giving the difference between the prediction and the actual

› Minimizing the error signal minimizes the data to be sent
  – Maximizes coding performance

› Generating a better prediction gives better coding performance
  – MPEG-4 AVC generates a better prediction than MPEG-2 Video

› Better coding of the error signal gives better coding performance
  – MPEG-4 AVC codes the error signal better than MPEG-2 Video
**GENERIC ENCODER BLOCK DIAGRAM**

- **Video Input** → **Pre-Processing** → **Temporal Model**
- **Previous/Future Frames** → **Spatial Model** → **Entropy Encoder**
- **Frame differences** → **Spatial Model**
- **Coefficients** → **Vectors**
- **Encoded Bit-stream**

- Removes temporal redundancy
- Removes spatial redundancy
- Lossless compression

Adapted from “H.264 and MPEG-4 Video Compression”, Iain E.G. Richardson
MPEG-4 AVC VS. MPEG-2 VIDEO: WHY/HOW IS IT BETTER?

Lower bit rates from technology advances and improved tools

› Technology → Moore’s Law
  - MPEG-2 standards developed over 10 years ago
  - ~1992: IC tech >1 micron, memory >$50/MB, processors <500MHz
  - ~2006: IC tech <0.1 micron, memory <$50/GB, processors >3GHz

› Improved tools → Similar overall signal processing flows – but with more and improved algorithmic tools
  - Signal processing stage: More parallel processing and more memory look-ups/stores, with selection of most bit-efficient result
  - Bit stream syntax: More efficient use of “overhead”/structure
  - Bit stream processing stage: More computationally intensive entropy coding
THE “PHILOSOPHY” OF ENGINEERING

› Computational complexity
  – A key factor in determining the technical success or failure of a codec

› What is “Quality”?
  – Impact of objective (rate-distortion curve) vs. subjective observations

› Tool and algorithm choices
  – Need to be taken as a whole, not individually
  – Optimizing for one content selection vs. being efficient for a wide range

› Bit-rates are lower
  – Traditional tool selection & optimization knowledge may not translate directly

› Theory vs. practice: Analogous issues apply
  – Differences in implementations may vary more than any inherent differences in the efficiencies of the algorithms themselves
MPEG-2 VS. MPEG-4 AVC OVERVIEW

› Intra Prediction Modes – fixed vs. many
› Inter Prediction Modes – some vs. many
› Motion Compensation – ½ pel vs. ¼ pel
› Transform – 8x8 vs. adaptive block size (4x4, 8x8)
› Transform – DCT vs. simpler integer
› Deblocking Filter – none vs. in-loop
› Entropy Coder – fixed vs. context adaptive
MPEG-2 ENCODER

Frame / (current)  

Motion Vector Prediction  

Motion Estimation  

Motion Compensaion  

Frame /-2 (reconstructed references)  

Transform  

Quantize  

Zig-zag Scan & Run Length Encode  

Entropy Coding  

reconstructed frame  

F  

E  

D  

Fm  

Fr  

Fr  

Er  

E  

Transform  

Inverse Transform  

Inverse Quantize  

inter coding  

predicted block  

intra coding  

motion vector  

+  

+  

-  

+  

-
MPEG-4 AVC REFERENCE ENCODER

- Frame n (current)
- Frame n-x (reconstructed references)
- Intra Prediction Mode Select
- Intra Prediction
- Motion Vector Prediction
- Motion Estimation
- Motion Compensation
- Transform
- Quantize
- Zig-zag Scan & Run Length Encode
- Entropy Coding

- In Loop Filter
- reconstructed frame
- + -
- Inter coding
- predicted block
- Inverse Transform
- Inverse Quantize

Zig-zag Scan & Run Length Encode
- + -
- Inter coding
- predicted block
- Inverse Transform
- Inverse Quantize

- + -
- Inter coding
- predicted block
- Inverse Transform
- Inverse Quantize

© Ericsson AB 2010
MPEG-2 vs. MPEG-4 AVC: Intra Prediction Modes

Block to Intra Code

MPEG-2 Modes

MPEG-4 AVC Modes

MPEG-2 Modes

DC Pred
Horz Pred
Vert Pred
Plane Pred

Prediction

Error
Signal
Encoded

© Ericsson AB 2010
MPEG-2: INTER PREDICTION MODES

Block to Predict

Reference Picture

16x16

16x8
MPEG-4 AVC: INTER PREDICTION MODES

Block to Predict

16x16

16x8

8x16

8x8

4x4

Reference Picture
Motion estimation is used to identify moving image content in order to better exploit temporal redundancy
MOTION ESTIMATION

Motion searching during encoding identifies the “best fit” between the current picture macroblock and the reference picture(s)

Source: H.264 and MPEG-4 Video Compression, Iain E. G. Richardson, Wiley
MOTION ESTIMATION

Only differences between pictures are encoded

Source: H.264 and MPEG-4 Video Compression, Iain E. G. Richardson, Wiley

Delta No ME

© Ericsson AB 2010
MOTION ESTIMATION

Source: H.264 and MPEG-4 Video Compression, Iain E. G. Richardson, Wiley

Delta 8x8 ME
MOTION ESTIMATION

Source: H.264 and MPEG-4 Video Compression, Iain E. G. Richardson, Wiley

Delta 4x4 ME
MOTION ESTIMATION

MPEG-4 AVC provides more ME block sizes

MPEG-2
16x16

MPEG-4 AVC
16x16
16x8, 8x16
8x8
8x4, 4x8
4x4

Source: H.264 and MPEG-4 Video Compression, Iain E. G. Richardson, Wiley

Appropriate block size choices
MPEG-2 VS. MPEG-4 AVC: QUARTER PEL MOTION COMPENSATION

Motion Comp Options

- Integer Pixels
- Half pel (MPEG-2)
- Quarter pel (MPEG-4 AVC)
MPEG-2 VS. MPEG-4 AVC: TRANSFORM

- Smaller block size available than MPEG-2 (4x4 vs. 8x8*) and adaptive block-size selection
- Simpler transform (computationally easier but same effect)
MPEG-2 VS. MPEG-4 AVC: IN-LOOP DEBLOCKING FILTER

Without In-Loop Filter (MPEG-2 Video)  

With In-Loop Filter (MPEG-4 AVC)
MPEG-2 VS. MPEG-4 AVC: ENTROPY ENCODER

› Removes statistical redundancy from bit stream

› MPEG-2: Variable Length Coding

› MPEG-4 AVC: CABAC (Context-based Adaptive Binary Arithmetic Coding)

› Context-based coding → Optimizes encoded bit string = reduces bits used
<table>
<thead>
<tr>
<th>Tool</th>
<th>Detailed Comparison of Tools (1 of 3)</th>
<th>Tool</th>
<th>Detailed Comparison of Tools (1 of 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intra Prediction</strong></td>
<td><strong>MPEG-2 Video (H.262)</strong></td>
<td><strong>MPEG-4 AVC (H.264)</strong></td>
<td></td>
</tr>
<tr>
<td>None: MB encoded</td>
<td>- DC predictors</td>
<td>- 4x4 Spatial</td>
<td></td>
</tr>
<tr>
<td>DC predictors</td>
<td></td>
<td>- 16x16 Spatial</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- I_PCM</td>
<td></td>
</tr>
<tr>
<td><strong>Picture Coding Type</strong></td>
<td><strong>MPEG-2 Video (H.262)</strong></td>
<td><strong>MPEG-4 AVC (H.264)</strong></td>
<td></td>
</tr>
<tr>
<td>- Frame</td>
<td>- Field</td>
<td>- Frame</td>
<td></td>
</tr>
<tr>
<td>- Field</td>
<td>- Picture AFF</td>
<td>- Field</td>
<td></td>
</tr>
<tr>
<td>- Picture AFF</td>
<td></td>
<td>- Picture AFF</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- MB AFF</td>
<td></td>
</tr>
<tr>
<td><strong>Motion Compensation Block Size</strong></td>
<td><strong>MPEG-2 Video (H.262)</strong></td>
<td><strong>MPEG-4 AVC (H.264)</strong></td>
<td></td>
</tr>
<tr>
<td>- 16x16</td>
<td>- 16x8, 8x16</td>
<td>- 16x16</td>
<td></td>
</tr>
<tr>
<td>- 16x8, 8x16</td>
<td></td>
<td>- 16x8, 8x16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 8x8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 8x4, 4x8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 4x4</td>
<td></td>
</tr>
</tbody>
</table>
# Detailed Comparison of Tools (2 of 3)

<table>
<thead>
<tr>
<th>Tool</th>
<th>MPEG-2 Video (H.262)</th>
<th>MPEG-4 AVC (H.264)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Motion Vector</strong></td>
<td>- Full Pel</td>
<td>- Full Pel</td>
</tr>
<tr>
<td><strong>Precision</strong></td>
<td>- Half Pel</td>
<td>- Half Pel</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Quarter Pel</td>
</tr>
<tr>
<td><strong>P Frame Feature</strong></td>
<td>- Single Reference</td>
<td>- Single Reference</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Multiple Reference</td>
</tr>
<tr>
<td><strong>B Frame Feature</strong></td>
<td>- 1 Reference Each Way</td>
<td>- 1 Reference Each Way</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Multiple Reference</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Direct &amp; Spatial</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Direct Modes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Weighted Prediction</td>
</tr>
</tbody>
</table>
### Detailed Comparison of Tools (3 of 3)

<table>
<thead>
<tr>
<th>Tool</th>
<th>MPEG-2 Video (H.262)</th>
<th>MPEG-4 AVC (H.264)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-Loop Filters</td>
<td>- None</td>
<td>- De-Blocking</td>
</tr>
<tr>
<td>Entropy Coding</td>
<td>- VLC</td>
<td>- CAVLC - CABAC</td>
</tr>
<tr>
<td>Transform</td>
<td>- 8x8 DCT</td>
<td>- 4x4 Integer “DCT” - 8x8 Integer “DCT”</td>
</tr>
<tr>
<td>Other</td>
<td>- Quantization Scaling Matrices</td>
<td>- Quantization Scaling Matrices</td>
</tr>
</tbody>
</table>
# MPEG-4 AVC – MAJOR ALGORITHMIC ENHANCEMENTS

<table>
<thead>
<tr>
<th>MPEG-2</th>
<th>MPEG-4 AVC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only one reference frame for P</td>
<td>Multiple reference frames</td>
</tr>
<tr>
<td>And two for B pictures</td>
<td>Bi-predictive modes for P and B pictures</td>
</tr>
<tr>
<td>Large square Motion Compensation blocks, 16x16</td>
<td>Various block shapes, down to 4x4</td>
</tr>
<tr>
<td>½ pel motion vector</td>
<td>¼ pel motion vector</td>
</tr>
<tr>
<td>8x8 DCT</td>
<td>4x4 integer transform</td>
</tr>
<tr>
<td>Single spatial prediction mode</td>
<td>Many intra prediction modes</td>
</tr>
<tr>
<td>Fixed quantization weighting</td>
<td>Integrated de-blocking filter</td>
</tr>
<tr>
<td>Variable Length Coding (Huffman)</td>
<td>Extended and finer control quantization</td>
</tr>
<tr>
<td></td>
<td>Context Adaptive VLC and Context Adaptive Binary Arithmetic Coding</td>
</tr>
</tbody>
</table>
MPEG-4 AVC / H.264 VERSION 2
FIDELITY RANGE EXTENSIONS ("FREXT")

› Version 1 primarily developed for direct-to-home (DTH) applications – broadcast, video discs, Internet

› Version 2 added Fidelity Range Extensions for professional applications
  – Contribution (front hauls), Distribution (back hauls)
  – Mastering, editing, archiving

› Extends color space and bit-depth (precision)
MPEG-4 AVC / H.264 EXTENSIONS: SCALABLE VIDEO CODING (SVC)

› Multi-layer encoding
   – The video bitstream contains one or more subset bitstreams that can each be decoded separately
   – Typically a Base Layer + one or more enhancement layers

› Useful for when a network or transmission medium has varying bandwidth

› Applications
   – Adaptive streaming
   – Tiered program offerings
   – Multi-device delivery (e.g., HDTV, PC, Mobile/Handheld)

› Profiles
   – Scalable Baseline (conversational, surveillance)
   – Scalable High (broadcast TV, streaming, storage)
   – Scalable High Intra (professional)

› Standard approved July 2007 (AVC Annex G)
   – Not yet widely deployed
SVC TYPES

› Temporal scalability (frame rate)
  – The video is encoded at multiple frame rates
  – The motion compensation dependencies are structured so that complete pictures can be dropped from the bitstream

› Spatial scalability (picture size)
  – The video is coded at multiple spatial resolutions
  – The data and decoded samples of lower resolutions can be used to predict data or samples of higher resolutions in order to reduce the bit-rate to code the higher resolutions

› SNR/Quality/Fidelity scalability
  – The video is coded at a single spatial resolution but at different qualities
  – The data and decoded samples of lower qualities can be used to predict data or samples of higher qualities in order to reduce the bit-rate to code the higher qualities
MPEG-4 AVC / H.264 EXTENSIONS: MULTIVIEW VIDEO CODING (MVC)

- Enables the efficient encoding of sequences captured simultaneously from multiple cameras using a single video stream

- Stereoscopic High Profile standardized in July 2009
  - First application: 3D content on Blu-ray Disc
  - Forward compatible with 2D decoders
    - Enhancement information is ignored and main layer can be decoded

- Other multiview Profiles are still work in progress

- The recent announcements regarding 3D broadcast TV do not use MVC
  - Most DTH providers initially will be using frame-compatible multiplexed pictures (1/2 horizontal resolution per eye)
AGENDA

› Overview of MPEG-4 AVC

› Transport considerations

› Issues to consider
  – Concatenation of decode/encode stages
  – Legacy receiver devices
  – Ancillary data & audio
  – Contribution
TRANSPORT CONSIDERATIONS

› Statistical multiplexing
› Alternative transport formats for IP-based delivery
› File delivery vs. streaming … bridging the gap
› Making streaming IP-based delivery more robust
STATISTICAL MULTIPLEXING AND MPEG-4 AVC

› Statistical multiplexing is used to spread bit-rate between services within a multiplex
  – Services which are instantaneously difficult to code get more bit-rate

› MPEG-4 AVC typically shows greater dynamic range in bit-rate use
  – Statistical multiplexing gives larger efficiency gains
  – More services per multiplex as compared to MPEG-2 HD
  – Statistical multiplexing efficiency improves with number of channels
THE MPEG-2 TRANSPORT STREAM

› Handles generic video, audio, data, metadata
  - Example video formats: MPEG-2 Video, MPEG-4 AVC, JPEG2k, …
› Maintains strict timing relationship among the constituent elementary streams
  - Clock forwarding of source content using 27MHz Program Clock References (PCRs) for exact receive-side reconstruction
  - Presentation times synchronized using 90kHz Presentation Time Stamps (PTSs)
MP4 FILE FORMAT

- Defined in ISO/IEC 14496-14
  - Based on the generic ISO File Format, ISO/IEC 14496-12
  - All data (video, audio, metadata, etc.) is encapsulated in boxes. Boxes contain all information needed to process
- Movie = the overall presentation
- Logically divided into tracks
- Each track represents a timed sequence of media (frames of video or audio)
  - Sample = timed unit within a track
  - Samples are implicitly numbered in sequence
  - A description defines how the sample may be decoded (which compression algorithm)
  - Includes description of codec type, needed codec resolutions, sampling rates, media durations
  - Timing information and synchronization between tracks defined in Sample Table Box (NTSC: TimeScale = 30000, sample duration = 1001)
THE CHALLENGES OF USING IP

› Broadcast demands
  – Intolerant of stream errors
    › Corrupt, lost or reordered packets
  – Stringent timing model
    › Need to synchronise source and receiver
    › MPEG-2 relies on a constant delay network
  – High data rates (2–50 Mbps)

› Network delivery
  › Provides best effort datagram delivery
    – Packets can be delivered out of order
    – Packets can be duplicated
    – Packets can be lost or discarded
    – Packets can be delayed for a long time

© Ericsson AB 2010
MAPPING MPEG-4 AVC INTO IP

› MPEG-2 TS is still required
  – TV set-top boxes are intrinsically TS aware
  – MPEG-2 TS provides packing and critical timing relationship of constituent video, audio, data and metadata streams
  – With typical stream bitrates involved for broadcast TV, the 2% TS packet header is not significant
    › Can become significant has bitrates drop – as in mobile video applications
MAPPING MPEG-4 AVC INTO IP

- Proposed mapping for new adaptive streaming techniques
- MPEG-2 TS no longer used … is it still needed?
- Real-time streaming: RTP timing to assume critical timing relationship among the constituent streams … will this work?

- Stateless streaming: “chunking” or media segmentation
  - CDN caches use HTTP to deliver files
  - Bridging/leveraging file-based work to do pseudo streaming techniques, like progressive download (PDL)
  - Next step: Adaptive streaming of live sessions
UDP/RTP MAPPING

MPEG-2 Transport Stream packets

IP | UDP | RTP | UDP payload

1 188 | 2 188 | 7 188

RTP layer adds timing information and sequence counter

IP Frame
RTP SOLVES TWO NETWORK PROBLEMS

Transmitted

Duplicate

Re-ordered

Decoded
RTP CAN’T SOLVE THIS

› Packets 4, 9 and 10 have been lost
› “IP” does not support retransmission, 4,9 and 10 are therefore gone forever
› Result is a corrupted stream
First FEC schemes for IP were proprietary
- COP3 is based on RTP Sequence Number
- RTP Frames are arranged in a rectangle of dimensions D * L
- FEC is added for every column and optionally for every row
- FEC payload is the same length as a single RTP payload (1-7 MPEG-2 TS Packets)
- FEC payload is the XOR of the row/column it protects
- Data carried over separate UDP ports
- Proven interoperability
- COP3 codified as SMPTE 2022
PACKET CORRECTION

Before:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>10</td>
<td>9</td>
<td>12</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
</tr>
</tbody>
</table>

After:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
</tr>
</tbody>
</table>

© Ericsson AB 2010
BURST CORRECTION NOT CORRECTED

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

© Ericsson AB 2010
ADD ROW FEC

and corrected
SMPTE 2022 FEC

› SMPTE 2022 Forward Error Correction is an open standard
  - Codified Pro-MPEG COP #3
› Proven interoperability of video equipment
  - VSF Interoperability demos in 2006 & 2007
› Increased robustness of connectivity
  - Designed specifically for IP circuits
› Delivers greater flexibility to users
  - Column only FEC mode for burst protection with minimum overhead
  - Column and Row FEC mode offering both burst and random protection
› Carriage of FEC packets external to transport stream
  - Ensures new and legacy receive devices remain operational
QUALITY TRADE OFFS

QUALITY

OVERHEAD

ROBUSTNESS

LATENCY
AGENDA

› Overview of MPEG-4 AVC
› Transport considerations
› Issues to consider
  – Concatenation of decode/encode stages
    – Legacy receiver devices
    – Ancillary data & audio
    – Contribution
TYPICAL CONTENT DELIVERY SCENARIO

- Pictures from SNG truck at e.g. sporting event
- National Network Headquarters
- Local Affiliate stations

**Creation**
- Satellite
- Telco

**Central or Regional**
- Satellite
- Telco

**Regional or Central**
- Mobile
- IPTV
- Cable
- Satellite
- Terrestrial

- Back-haul high bit-rate encoding (e.g., MPEG-2 @ ~50+ Mbps)
- Front-haul mid bit-rate encoding (e.g., MPEG-2 @ ~20-40 Mbps)
- Emissions/DTH low bit-rate encoding (e.g., MPEG-2 at ~15-18 Mbps)
Concatenation stages

- Goal: Preserve as much original picture quality as possible through multiple stages of encode/decode ➔ “Concatenation” impact
COMPARING CONCATENATION

HD “station” encoder at 15Mbps MPEG-2 Video emissions bit-rate

Graph shows the effects of different distribution bit-rates for MPEG-2 and MPEG-4 AVC, decoded then recoded to MPEG-2 emissions bit-rate

Relative PSNR vs. only final encode
Focusing on the likely operating region

- Example: ~24Mbps MPEG-2
  - Equivalent MPEG-4 AVC rate would be around 18Mbps, ~25% lower bit-rate
  - Generally, this will increase to some degree with improved MPEG-4 AVC compression performance

- Why isn’t the difference greater?
  - MPEG-2 to MPEG-2 always has some concatenation benefits at the macroblock level
    - The transform & quant matrices giving similar characteristics at the two stages
    - Macroblock structures & vector consistency
  - MPEG-4 AVC to MPEG-2 doesn’t offer any “re-use” of the artifacts
    - Transforms and quants are different from MPEG-2; therefore, artifacts are additive
CONCATENATION EXAMPLE: “BASKETBALL”

› “Front haul” stage
  – MPEG-2 @ 40Mbps vs. MPEG-4 AVC @25Mbps

› Emissions stage
  – MPEG-2 @ 16Mbps

› All formats 1920x1080i @29.97
FRONT HAUL: MPEG-2 40MBPS
FRONT HAUL: MPEG-4 AVC 25MBPS
MPEG-2 40MBPS -> MPEG-2 16MBPS
MPEG-4 AVC 25MBPS -> MPEG-2 16MBPS
FRONT HAUL: ZOOM IN

MPEG-2 40Mbps
› More noise
› More blocky
› More detail

MPEG-4 AVC 25Mbps
› Less noise
› Less blocky
› Less detail
› 60% of bit-rate

Which looks better to you? ...
EMISSIONS STAGE: ZOOM IN

MPEG-2 40Mbps → MPEG-2 16Mbps

MPEG-4 AVC 25Mbps → MPEG-2 16Mbps

Which looks better to you?

Doesn’t matter …
emissions stage dominates artifacts!

But there are still subtle differences
SUMMARY

› Comparison of MPEG-4 AVC vs. MPEG-2 concatenated into MPEG-2 at emissions/DTH bit-rates
  – Performance is dominated by the final MPEG-2 encode stage when bit-rate is significantly lower than previous stage
  – MPEG-4 AVC distribution …
    › provides better video performance than MPEG-2 Video at the same rate, or
    › provides approximately 25% bit-rate saving for the same re-encoded MPEG-2 quality
  – The difference is significant, but not as great as just MPEG-4 AVC compared to MPEG-2 individually
AGENDA

› Overview of MPEG-4 AVC
› Transport considerations
› Issues to consider
  – Concatenation of decode/encode stages
  – Legacy receiver devices
  – Ancillary data & audio
  – Contribution
THE LEGACY MPEG-2 SITUATION

› Broadcaster & Programmer content is fed to many cable systems
› Most cable systems are based on MPEG-2 Video infrastructure and set-top boxes
› There is often an older analog tier as well
MPEG-4 AVC DISTRIBUTION TO MPEG-2 NETWORK INTERFACING

Interfacing the MPEG-4 AVC broadcast to the MPEG-2 cable headends achievable through two methods

- Decode, re-encode
  - Expensive
  - Programmer has little control over final video quality
- Or transcode
TRANSCODING RECEIVERS

› New generation of receivers transcode MPEG-4 AVC to MPEG-2 Video
› Takes advantage of “golden combination” of MPEG-4 AVC and new modulation techniques, while being transparent to existing station / headend
› Economically viable
  – Other than replacing downlink receiver, no changes to existing headend equipment required
    › No new signal routers/switchers
    › No new MPEG-2 encoders
› When consumer STB deployment completes migration from MPEG-2 to combo MPEG-2 / MPEG-4 AVC, transcoding receivers can pass-thru MPEG-4 AVC as is
  – Future proof for service provider
SOLUTION EXAMPLE

Programmer Transcoder

Joint Descrambler plus service list

Transcode
Transcode
Transcode
Transcode
Transcode
Mux

Security

Serv 41 Serv 42 Serv 43 Serv 44 Serv 45 Serv 46
Serv 51 Serv 52 Serv 53 Serv 54 Serv 55 Serv 56

Serv 41 Serv 42 Serv 43 Serv 44 Serv 45 Serv 46
Serv 51 Serv 52 Serv 53 Serv 54 Serv 55 Serv 56

Entitled services decrypted
All services listed to Operator

Output MPEG-2 HD services

Tuner1 Demod 1
Tuner2 Demod 2

RF1
RF2

MPEG-4 AVC encrypted (likely stat muxed)
AGENDA

› Overview of MPEG-4 AVC

› Transport considerations

› Issues to consider
  – Concatenation of decode/encode stages
  – Legacy receiver devices
  – Ancillary data & audio
  – Contribution
ANCILLARY DATA HANDLING

› Closed captions
  - HD→SD conversion (CEA-708 -> CEA-608)
  - Carriage: MPEG-2 picture user_data -> AVC user SEI message

› Time code (VITC, LTC)
  - SMPTE 12M-2 time code in VANC
  - SMPTE 2038 carriage

› Active Format Description (AFD)
  - HD→SD down-conversion: change aspect ratio
  - SMPTE 2016 AFD in VANC
  - Carriage: MPEG-2 picture user_data -> AVC user SEI message

› SD VBI legacy – Nielsen AMOL, TVGuide, NABTS
  - SCTE 127 carriage, SMPTE 2031 (SCTE 127 in VANC)

› Private ancillary data carriage
  - SMPTE 2038 carriage
WHAT ABOUT AUDIO?

› For many years, digital broadcast channels have handled audio in the following formats
  - MPEG-1 Layer II
  - Linear PCM
  - Dolby E
  - Dolby Digital (AC-3) (2/0 and 5.1)
  - Digital Theater Sound (DTS)

› Newer audio formats have arisen as well
  - Advanced Audio Coding (AAC-LC and HE-AAC v1/v2)
  - Dolby Digital plus+
  - Dolby Pulse
**DON’T FORGET AUDIO BIT-RATE PLANNING!**

<table>
<thead>
<tr>
<th># of Audio Channels</th>
<th>MPEG Layer II</th>
<th>Dolby Digital (AC-3)</th>
<th>Dolby Digital Plus+</th>
<th>MPEG AAC-LC</th>
<th>MPEG HE-AAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stereo</td>
<td>192kbps</td>
<td>192kbps</td>
<td>96kbps</td>
<td>96kbps</td>
<td>64kbps (v1) 48kbps (v2)</td>
</tr>
<tr>
<td>5.1</td>
<td>Not Supported</td>
<td>384kbps - 448kbps</td>
<td>&lt;320kbps</td>
<td>256kbps</td>
<td>160kbps</td>
</tr>
</tbody>
</table>

*Claimed rates for CD quality for average-complexity content*

- Audio bit rates primary planning factor as video rates decrease
- Multiple languages per video service
- CBR – can’t be stat muxed

*For AVC SD, audio bit-rate allocation could be higher than video!*
AUDIO

› Many legacy devices do not decode the newer audio formats

› Either pass-through existing format at the higher bit-rate required

› Or, transcode from a more bit-rate efficient format to a legacy format

› Transcoding is possible in today’s DSPs
AGENDA

› Overview of MPEG-4 AVC
› Transport considerations
› Issues to consider
  – Concatenation of decode/encode stages
  – Legacy receiver devices
  – Ancillary data & audio
  – Contribution
CONTRIBUTION APPLICATION

› Global contribution
› High-end distribution
› Link
   – Satellite (DVB-S2), or
   – Fiber (IP/Ethernet)
### Differences between Contribution and DTH

#### System-level differences

<table>
<thead>
<tr>
<th>Contribution</th>
<th>DTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low latency</td>
<td>Low bandwidth</td>
</tr>
<tr>
<td>Best quality in given time</td>
<td>Best quality in given bandwidth</td>
</tr>
<tr>
<td>High bit-rate operations</td>
<td>Low bit-rate operations</td>
</tr>
<tr>
<td>First of multiple stream generations (4:2:2 mostly)</td>
<td>Last of multiple generations (4:2:0)</td>
</tr>
<tr>
<td>Small/medium infrastructure (flexibility)</td>
<td>Large/huge infrastructure (legacy + interoperability)</td>
</tr>
<tr>
<td>CBR</td>
<td>VBR</td>
</tr>
<tr>
<td>Device control (mostly)</td>
<td>System control</td>
</tr>
</tbody>
</table>

#### Encoder-level differences

<table>
<thead>
<tr>
<th>Contribution</th>
<th>DTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low latency</td>
<td>Statistical multiplexing</td>
</tr>
<tr>
<td>4:2:2</td>
<td>4:2:0</td>
</tr>
<tr>
<td>High bit-rate operations</td>
<td>Low bit-rate operations</td>
</tr>
<tr>
<td>CBR</td>
<td>VBR</td>
</tr>
<tr>
<td>Professional audio support (Dolby E, LPCM)</td>
<td>Consumer audio support (Layer II, AC-3, AAC)</td>
</tr>
<tr>
<td>Generic VBI</td>
<td>Specific VBI</td>
</tr>
<tr>
<td>Front panel + Web browser control of single unit</td>
<td>System level control and monitoring</td>
</tr>
</tbody>
</table>
MOVING CONTRIBUTION TO MPEG-4 AVC

› Need to see real benefits over MPEG-2
  – Bandwidth savings
  – Increased quality

› Need to build an infrastructure
  – End-to-end system
  – Interoperable equipment
  – Future proof solutions
NEW CONTRIBUTION SOLUTION

› MPEG-4 AVC HD 4:2:2
› Support for 10-bit precision operations
› Support for 1080p50/60
› Deterministic low latency
› High bit-rate
› IP/Ethernet, ASI, and/or DVB-S2 outputs
› Contribution audio support
  – Dolby E and LPCM
BENEFITS OF MPEG-4 AVC 4:2:2

MPEG-2 HD 4:2:2 @ 60 Mbps

MPEG-4 AVC HD 4:2:2 @ 40 Mbps

-30%
BENEFITS OF 10-BIT

MPEG-4 AVC HD 4:2:2 8-bit

MPEG-4 AVC HD 4:2:2 10-bit
BENEFITS OF 10-BIT

Subjective and objective improvement on 'low frequency' areas

High frequency source hides the benefit of 10-bit over 8-bit

But no penalty in using 10-bit, so no harm using for all content