Introduction to H.264 Video Codec Standard: Hardware Challenges and Opportunities

Ahmed Abu-Hajar, Ph.D
DIGITAVID, Inc
San Jose, CA
abuhajar@digitavid.net
(408) 506-2776
PO Box: 720998
San Jose, CA 95172
Presentation Outline

- Introduction to Video Processing
- Overview of H.264-AVC Standard
- Overview of Hardware Implementation
- Challenges and Opportunities
Overview of Video Processing

- Objectives
  - Getting familiar with digital images and digital video sequences
  - Getting familiar with common acronyms and terms commonly used among video engineers

Ahmed Abu-Hajar, Ph.D.
Digitavid, Inc
Overview of Video Processing

Q) What Is Digital Image?
A) Digital image is a 2-D signal which is composed of very small picture elements called pixels

Q) What Is Pixel?
A) Pixel is very small rectangular area which has a uniform intensity value

Ahmed Abu-Hajar, Ph.D.
Digitavid, Inc
Overview of Video Processing

Q) How Images Display Colors?
A) Each pixel has three primary colors:
   - Red
   - Green
   - Blue

Note:
   - Human’s vision is less sensitive to Red and Blue
   - In Video the colored pixels are down sampled for every 2x2 pixels into:
     - Two green samples called Luminance
     - One Red sample called Chrominance CR
     - One blue sample chrominance CB
Overview of Video Processing

Q) What is Video?

A) Video is a sequence of images
   - Each image is called frame
   - Each frame has equal displaying time
     \[ \Delta t = \frac{1}{30} \text{s} \text{ or } \Delta t = \frac{1}{25} \text{s} \]
Overview of Video Processing

More on Frames

- Each frame is decomposed into two fields
  - **Even Field**: Includes the even rows of the frame
  - **Odd Field**: Includes the odd rows of the frame
- Two types of fields
  - **Progressive Scanned Field**: The even and odd fields are captured at the same time
  - **Interlaced Scanned Field**: The even and the odd fields are captured at different times
Overview of Video Processing

Segmenting Each Frame into Blocks

Macroblock (MB)

- Each Frame is segmented into big blocks called macroblocks
- Each macroblock contains:
  - One 16×16 Y component
  - One 8×8 CB component
  - One 8×8 CR component
Overview of Video Processing

Basics About Compression

- More probable values require less bits (Entropy Coding)
- Frames (or images) are smooth (low frequency)
  - Transform the frame into frequency domain (or DCT-Domain)
    - Most of the information energy is concentrated in FEW COEFFICIENTS
    - LOW ENERGY coefficients are truncated to ZEROS
    - The more we truncate to zero, the lower the quality of the video
    - Only few coefficients needed to be coded ➔ Compression
Overview of Video Processing

More on Compression

- The **Objective** is to create many zeros without lowering the video quality
- Prediction is used to predict the current pixel value.
- Prediction reduces Quantization (truncation) error
- Two types of predictions
  - **Intra Prediction**: Predicts current pixel value from the current frame
  - **Inter prediction**: Predicts current pixel value from previously coded frames
Introduction to H.264-AVC Standard

Scope of the Standard

The main objective is to efficiently compress video sequences

Ahmed Abu-Hajar, Ph.D.
Digitavid, Inc
Introduction to H.264-AVC Standard

Coder Structure
- Contains two layers
  - Video Coding Layer (VCL)  
    - Compress video content
  - Network Abstract Layer (NAL)  
    - Organize data to according to networks protocols

Ahmed Abu-Hajar, Ph.D.
Digitavid, Inc
Introduction to H.264-AVC Standard

Video Coding Layer (VCL)

- **Two Coding Modes**
  - **Intra Coding**: Does not use previously coded frames
    - Segment current frame into blocks
    - For each block: Predict pixels values from previously coded pixel
    - Transform the predicted block
    - Quantize the coefficients
    - Entropy code the quantized coefficients
  - **Inter Coding**: Uses previously coded frames to predict current frame
    - Segments the current frame into blocks
    - For each block in the current frame, search for the best block from previously coded frames that minimizes the error. (Motion Estimation)
Introduction to H.264-AVC Standard

Block Diagram

- Intra Prediction
- Transform
- Quantize Coefficients
- Intra Coding
- Entropy Coding
- Compressed Data
- Motion Compensation
- Inter Coding
- Inverse Quantization
- Inverse Transform
- Motion Estimation Algorithm
- De-blocking Filter & Inverse Prediction
- Current Block
Introduction to H.264-AVC Standard

Block Diagram

- Intra Prediction
- Transform
- Quantize Coefficients
- Entropy Coding
- Compressed Data

Intra Coding

- Inverse Quantization
- Inverse Transform

Inter Coding

- Motion Compensation
- Motion Estimation Algorithm
- De-blocking Filter & Inverse Prediction
Introduction to H.264-AVC Standard

- **Frame Format**
  - Monochrome – contains one color
  - Colored frame (Y, Cr, Cb)
    - 4:4:4 ➔ same size for (Y, Cr, Cb)
    - 4:2:2 ➔ Y is full size, (Cr, Cb) same height half width
    - 4:2:0 ➔ Y is full size, (Cr, Cb) half height and half width

- Each frame is segmented into Macroblocks
  - Y component (16x16)
  - Cb and Cr components (8x8) blocks

- Each macroblock is segmented into (16x8), (8x16), (8x8), (8x4), (4x8) and (4x4)
Introduction to H.264-AVC Standard

Block Diagram
Introduction to H.264-AVC Standard

- **Intra Prediction Modes:**
  - Lumina and chromina components are different
    - I_PCM: Skip prediction and transform
    - 16x16 Prediction Mode
      - Vertical, Horizontal, DC or Plane
      - Select the best mode
    - 4x4 Prediction Mode
      - Vertical, Horizontal, DC, Diagonal Down Left, diagonal Down Right, Vertical Right, Horizontal down, Vertical Left, Horizontal Up
- Coder must select the prediction mode with the least bit rate
- Each mode uses addition and shift operation
Introduction to H.264-AVC Standard

- Current Block
- Prediction Mode 1
- Prediction Mode 2
- Prediction Mode N

Select Best Prediction

Predicted Current Block
Introduction to H.264-AVC Standard

Block Diagram

- Intra Prediction
- Transform
- Quantize Coefficients
- Entropy Coding
- Compressed Data

- Intra Coding

- Inter Coding
  - Motion Compensation
  - Motion Estimation Algorithm
  - De-blocking Filter & Inverse Prediction
  - Inverse Quantization
  - Inverse Transform
Introduction to H.264-AVC Standard

Block Transform

- Compact the energy of the coefficients into the low frequency components
- Applied to 4x4 blocks
- Use integer Transforms (to reduce Hardware Complexity)

\[
\begin{bmatrix}
c_{00} & c_{01} & c_{02} & c_{03} \\
c_{10} & c_{11} & c_{12} & c_{13} \\
c_{20} & c_{21} & c_{22} & c_{23} \\
c_{30} & c_{31} & c_{32} & c_{33} \\
\end{bmatrix}
= \begin{bmatrix}
1 & 1 & 1 & 1 \\
2 & 1 & -1 & -2 \\
1 & -1 & -1 & 1 \\
1 & -2 & 2 & -1 \\
\end{bmatrix}
\begin{bmatrix}
p_{00} & p_{01} & p_{02} & p_{03} \\
p_{10} & p_{11} & p_{12} & p_{13} \\
p_{20} & p_{21} & p_{22} & p_{23} \\
p_{30} & p_{31} & p_{32} & p_{33} \\
\end{bmatrix}
= \begin{bmatrix}
1 & 2 & 1 & 1 \\
1 & 1 & -1 & -2 \\
1 & -1 & -1 & 2 \\
1 & -2 & 1 & -1 \\
\end{bmatrix}
\]

Ahmed Abu-Hajar, Ph.D.
Digitavid, Inc
Introduction to H.264-AVC Standard

Quantize Coefficients
- One of 52 (QP) Quantizers is selected
- Requires Look-Up Tables

Entropy Coding
- Selects one of Two Coders
  - Context-Adaptive Variable Length Coding (CAVLC) (Superior but Complex)
  - Exp-Golomb code with regular decoding properties (Simple but Inferior)
Introduction to H.264-AVC Standard

Block Diagram

Intra Prediction -> Transform -> Quantize Coefficients

Intra Coding

Entropy Coding -> Compressed Data

Inter Coding

Motion Compensation

Inverse Quantization

Inverse Transform

De-blocking Filter & Inverse Prediction

Motion Estimation Algorithm

Quantize Coefficients
Introduction to H.264-AVC Standard

Block Diagram

- Intra Coding
  - Intra Prediction
  - Transform
  - Quantize Coefficients
  - Entropy Coding
  - Compressed Data

- Inter Coding
  - Motion Compensation
  - Motion Estimation Algorithm
  - De-blocking Filter & Inverse Prediction
  - Inverse Quantization
  - Inverse Transform
Introduction to H.264-AVC Standard

Block Diagram

Intra Coding
- Intra Prediction
- Transform
- Quantize Coefficients
- Entropy Coding
- Compressed Data

Inter Coding
- Motion Compensation
- Inter Coding
- Inverse Quantization
- Inverse Transform
- De-blocking Filter & Inverse Prediction
- Motion Estimation Algorithm
Introduction to H.264-AVC Standard

Inter Coding
- Predicts the current block from previously coded frames
- The process is called Motion Estimation (ME)
- The block moves due to the moving object within the scene
Introduction to H.264-AVC Standard

Inter Coding

- Motion search algorithms
  - Full search
  - Three step search
  - Diamond search
  - Many more
Introduction to H.264-AVC Standard

Inter Coding

- Search within the most four previous frames

Ahmed Abu-Hajar, Ph.D.
Digitavid, Inc
Introduction to H.264-AVC Standard

Q) How is the motion Estimation Vector is selected?

- A) Based on the sum of Absolute Difference between two blocks

\[
SAD(u, v) = \sum_{x=1}^{N} \sum_{y=1}^{M} \left| p_{k,(r,s)}(x, y) - p_{k+1,(u,v)}(x, y) \right|
\]

- \( p_{k,(r,s)} \) is the pixel value in the \( k^{th} \) frame within block \( (r,s) \).
- \( p_{k+1,(u,v)} \) is the pixel value in the \( (k+1)^{th} \) frame within block \( (u,v) \).

The Block with the least SAD is Selected

Ahmed Abu-Hajar, Ph.D.
Digitavid, Inc
Introduction to H.264-AVC Standard

Q) Motion Estimation is Very COMPLEX, why?
- Must calculate SAD for each candidate block
- It compares the best candidate block
- H.264 uses variable block-size for motion estimation
- H.264 interpolate values between samples for (1/8) pixel resolution
Hardware Overview

- **Top-down Design Methodology**
  - **Specifications**
    - Set by the standard
  - **Subsystem**
    - Break the system into modules
  - **Register**
    - Implement subsystem using registers
  - **Gate**
    - Implement registers using gates
  - **Transistor**
  - **Layout**
  - **Fabrications**
  - **Testing and Verifications**
Hardware Overview

- **Most common hardware architectures**
  - **Single Instruction Multiple Data (SIMD)**
    - One Control Unit Multiple processing units
    - Load data into storage elements
    - Efficient for microprocessors design
  - **Systolic**
    - Pipelined structure
    - Each processing unit has its own control unit
    - More flexible
    - Less interconnects
    - Synchronized data flow through I/O
Hardware Overview

- Mesh Type SIMD array

![Diagram of mesh type SIMD array with control unit, control bus, data bus, and interconnection network.](image)
Hardware Overview

- Systolic Array.

![Diagram of hardware overview with control units, processing units, and an interconnection network.]

Ahmed Abu-Hajar, Ph.D.
Digitavid, Inc
Hardware Overview

- Types of Systolic Structures

  - Linear array with 1D I/O. This configuration is suitable for single I/O.

  - Linear array with 2D I/O. It allows more control over linear array.

Ahmed Abu-Hajar, Ph.D.
Digitavid, Inc
Hardware Overview

- **Types of Systolic Structures**
  - **Planar array** with perimeter I/O. This configuration allows I/O only through its boundary cells.

*Other Structures do exist*
Example of Hardware Architecture

HDTV  IEEE Transaction on Video Technology June 2006

Ahmed Abu-Hajar, Ph.D.
Digitavid, Inc
Hardware Overview

- Used four pipeline stages
- Intensive local SRAM units “Caches” requirements for each stage
- High level of parallelism for each stage “Tree structure”
Summary

- We briefly presented basics of image processing
- We briefly presented H.264-AVC Structure
- We briefly presented basics of systolic Hardware design
- We briefly presented most recent HDTV H.264-AVC chip