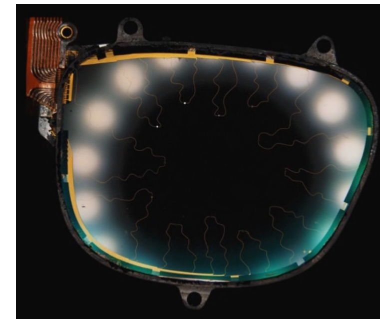
Meta Connect:
Orion AR Glasses

Tiny Projectors

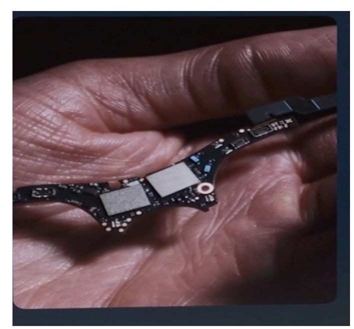
Light Waveguides



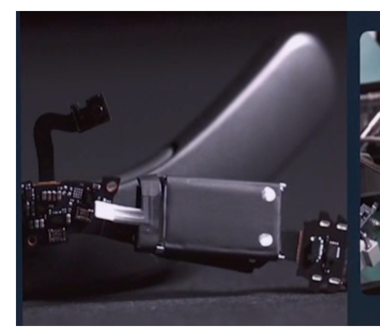
Nanoscale
3D Structures

Holograms

holographic displays, with eye tracking, hand tracking, and a neural-interface wristband, Lightweight (all under 100gms)



Custom Sensors



Battery

EECE5698
Networked XR Systems

Lecture Outline for Today

- Network Capacity vs. Requirements of Applications
- Compression Fundamentals
- 2D Video Compression

Today's Internet

- Wired
  - Fiber, Cable
- Wireless
  - Cellular
  - WiFi
  - Satellite

Internet speeds

- What are the max speeds for today's Internet?
  - Wired
  - Cellular
  - WiFi
  - Satellite

Internet type	Max speed
Fiber	10,000Mbps (5 Gbps)
Cable	1,200Mbps (1.2 Gbps)
DSL	100Mbps
5G	1,000Mbps (1 Gbps)
4G LTE	9–50Mbps
Fixed wireless	100Mbps
Satellite	100Mbps

https://www.highspeedinternet.com/how-much-internet-speed-do-i-need

What are the average speeds?

- Wired = ~1gbps
- WiFi = ~ 100Mbps
- Cellular = ~ 100Mbps

- Depends on the location
  - Campuses, Homes, Urban, Rural, Country (Developed vs. Developing worlds)
  - Many factors

How much Internet speed you need?

	Minimum	Recommended
Email	1Mbps	1Mbps
Web browsing	3Mbps	5Mbps
Social media	3Mbps	10Mbps
Streaming SD video	3Mbps	10Mbps
Streaming HD video	5Mbps	25Mbps
Streaming 4K video	25Mbps	100Mbps
Online gaming	5Mbps	100Mbps
Streaming music	1Mbps	5Mbps
One-on-one video calls	1Mbps	25Mbps
Video conference calls	2Mbps	50Mbps

https://www.highspeedinternet.com/how-much-internet-speed-do-i-need

2D Video as an Example

- How much bandwidth does a 2D movie needs
  - Example: 2-hour movie, 30 Fps, 8-bit depth, 1080p
  - Total = 2x60x60x30x3x1920x1080 Bytes
    or 1.25TB or 1.4Gbps
  - On a home WiFi with say average 150Mbps speed, it takes about 19 hours to download this movie

2D Video as an Example

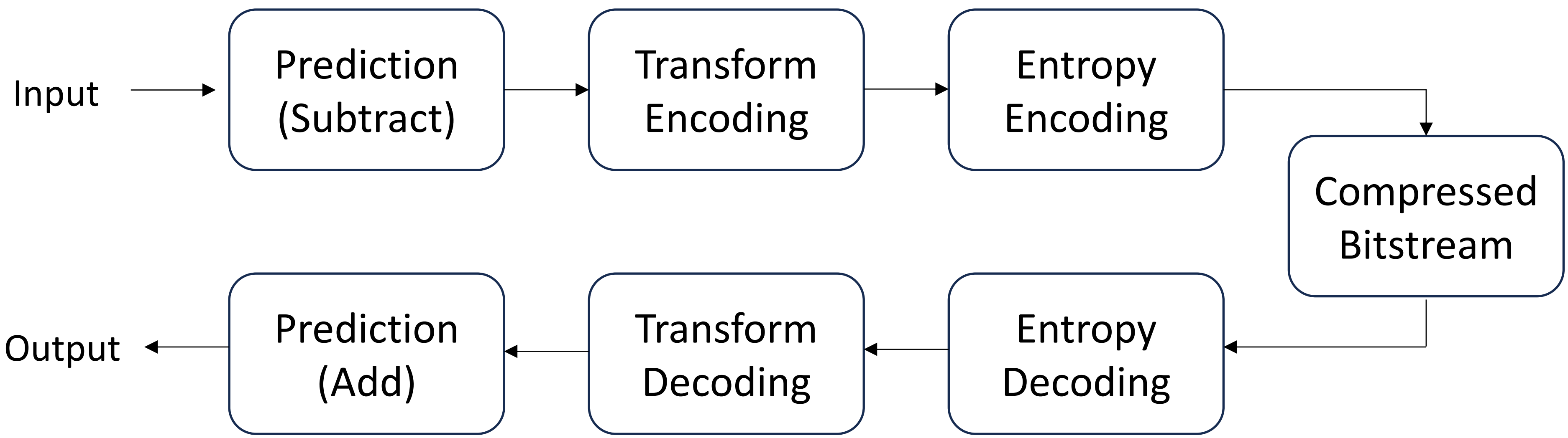
- But you’re watching your Netflix movie in real-time

Compression Fundamentals

- Two types of compression methods
  - Lossless
    - No loss of information
  - Lossy
    - There is some information loss.. But perceptually not much
    - Useful in case of poor Internet speeds

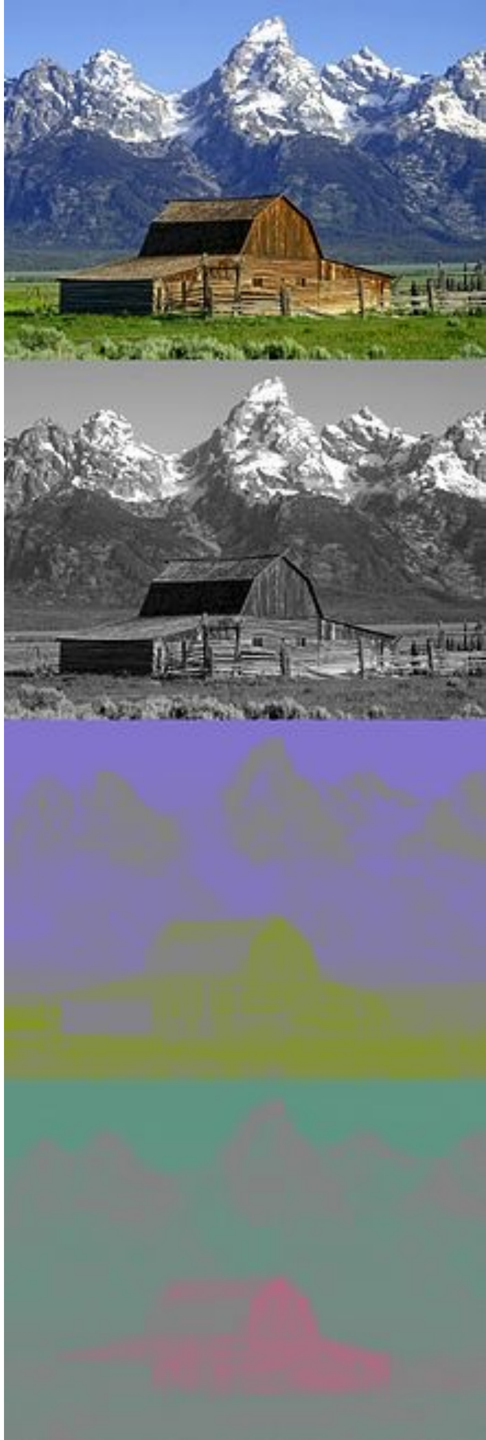
Compression Fundamentals

- Key steps involved in video compression pipeline
  - Color space or Chroma sub-sampling



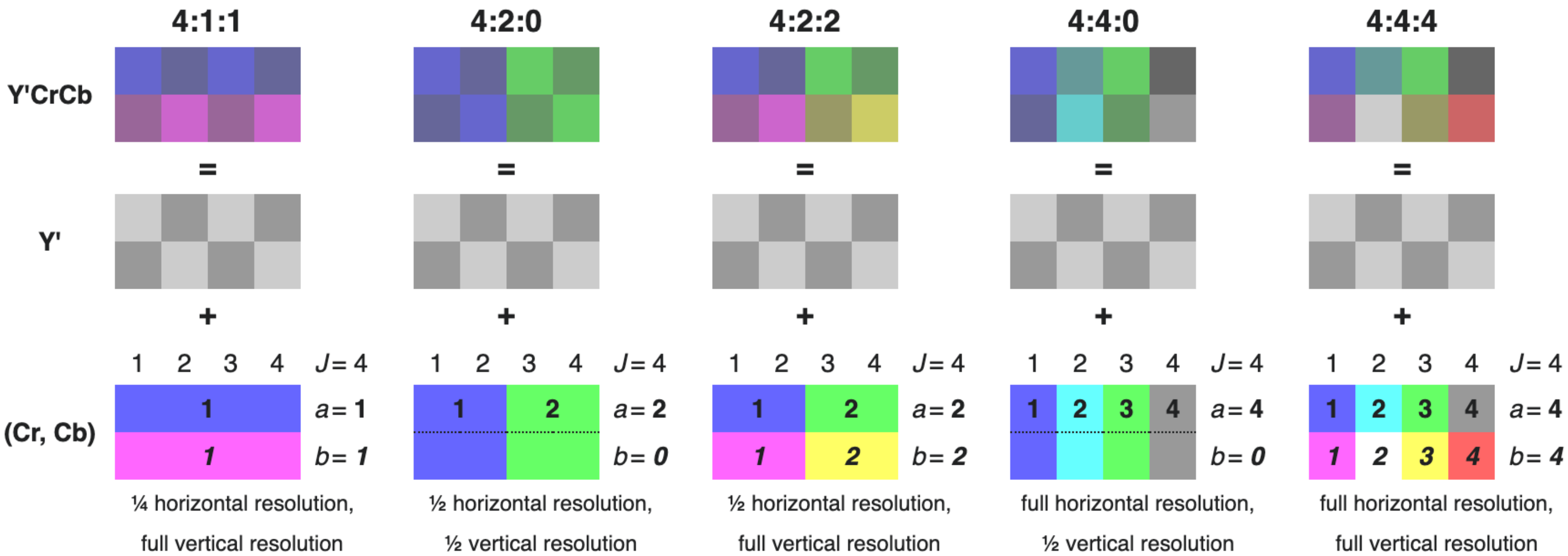
Chroma Sub-sampling

- RGB – 3 channels
  - Gives equal importance to all 3 channels
- YCbCr – 3 channels
  - Gives more importance to Luma
  - Less importance to Chroma
  - Perceptually minimal or no loss
- The Y image on the right is essentially a greyscale copy of the main image.



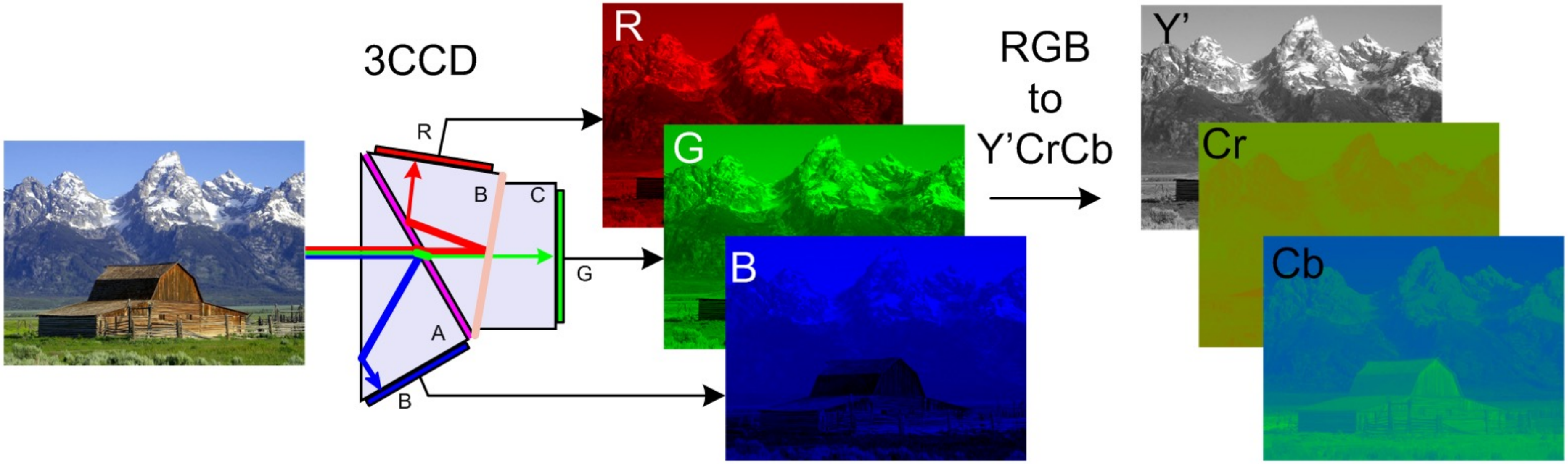
https://en.wikipedia.org/wiki/Chroma_subsampling

Chroma Sub-sampling



https://en.wikipedia.org/wiki/Chroma_subsampling

Chroma Sub-sampling



$$Y' = 16 + \frac{65.738 \cdot R'_D}{256} + \frac{129.057 \cdot G'_D}{256} + \frac{25.064 \cdot B'_D}{256}$$

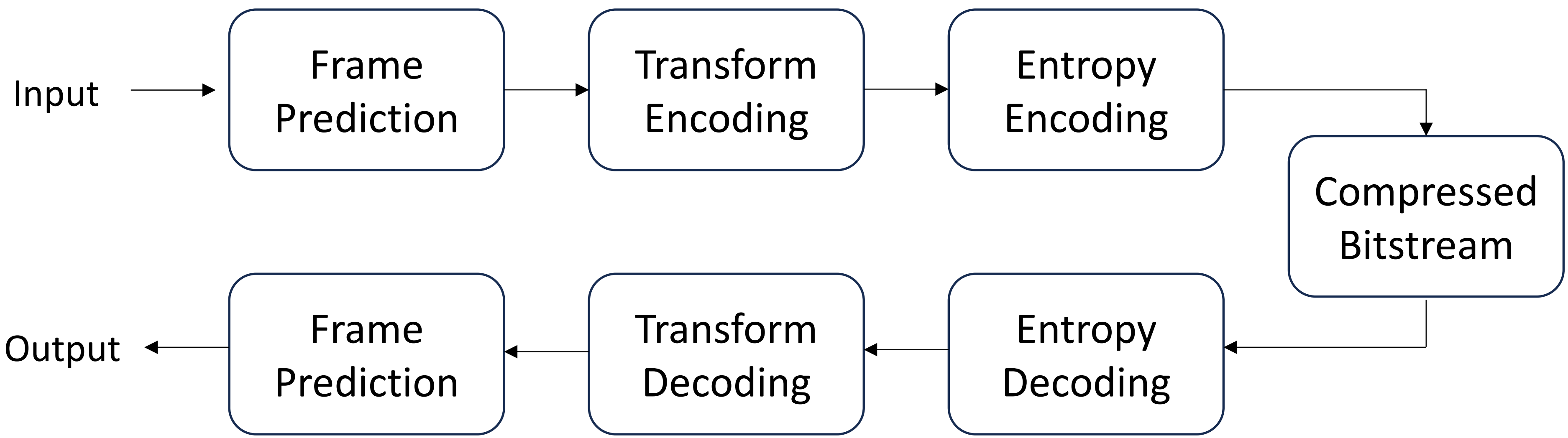
$$C_B = 128 - \frac{37.945 \cdot R'_D}{256} - \frac{74.494 \cdot G'_D}{256} + \frac{112.439 \cdot B'_D}{256}$$

$$C_R = 128 + \frac{112.439 \cdot R'_D}{256} - \frac{94.154 \cdot G'_D}{256} - \frac{18.285 \cdot B'_D}{256}$$

https://en.wikipedia.org/wiki/Chroma_subsampling

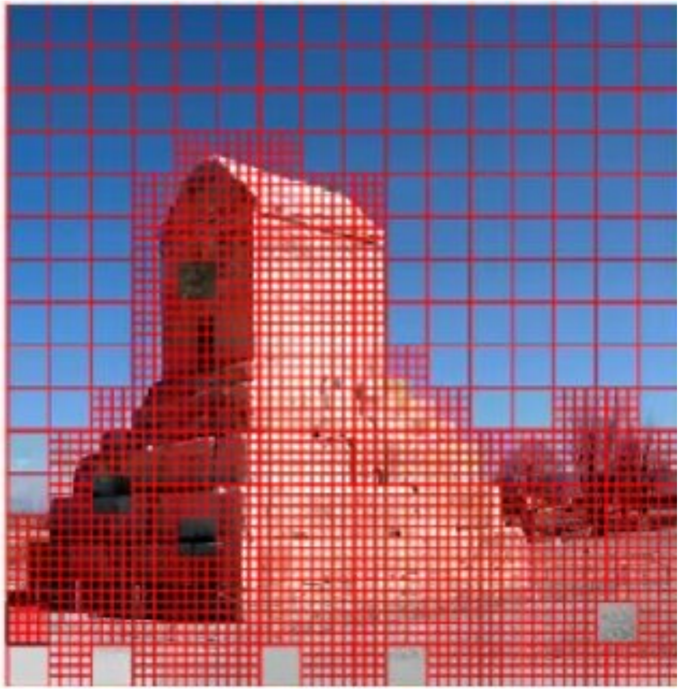
Compression Fundamentals

- Key steps involved in video compression pipeline
  - Color space or Chroma sub-sampling



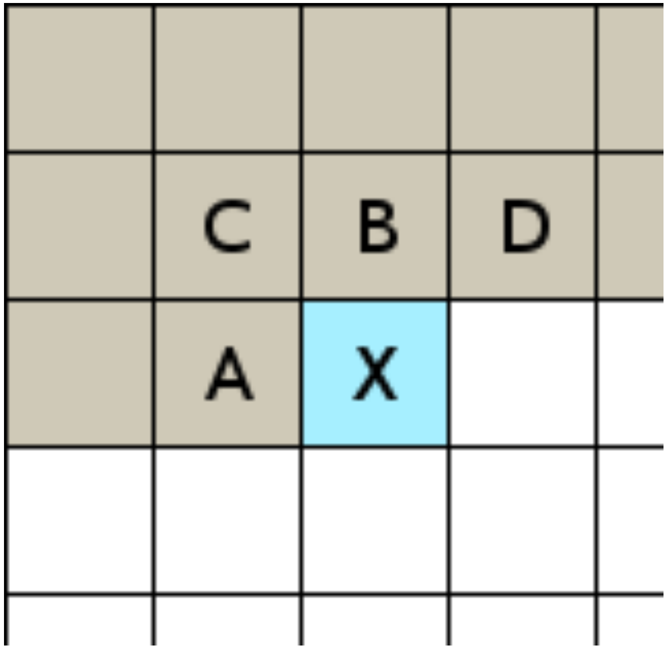
Frame Prediction

- Exploiting redundancy in the video content
  - Intra frame prediction
    - Within the frame – spatial redundancy
  - Inter frame prediction
    - Across the frames– temporal redundancy



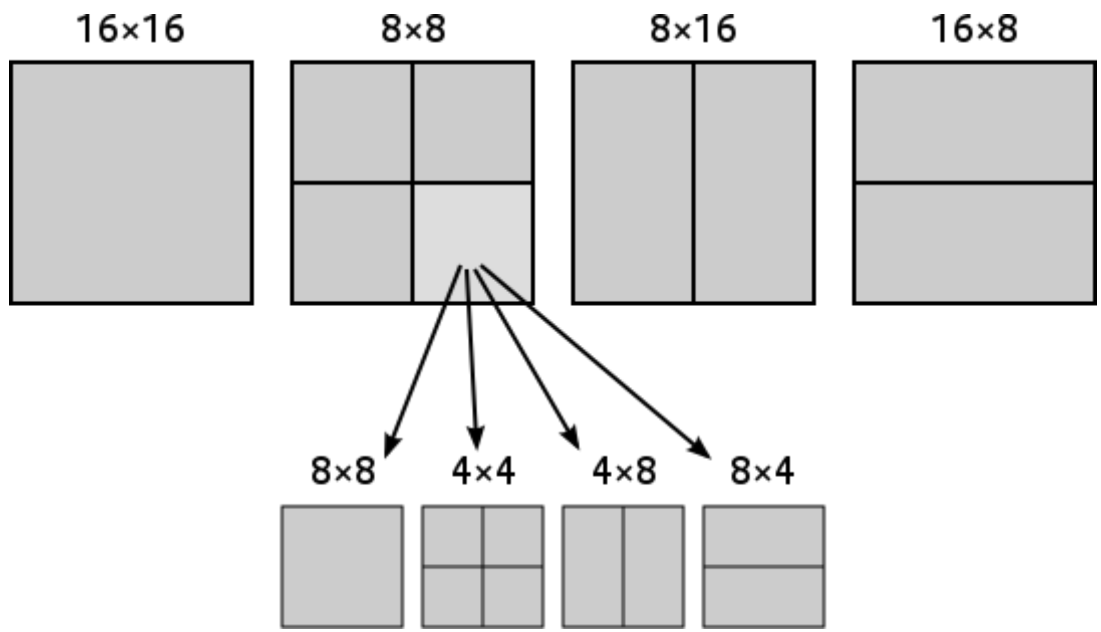
Frame Prediction

- Intra prediction
  - Since neighboring pixels within an image are often very similar, rather than storing each pixel independently, the frame image is divided into blocks and typically minor difference between each pixel can be encoded using fewer bits.



Frame Prediction

- Typical Block Sizes or Macroblock sizes



Latest compression algorithms can do up to 64x64 blocks of pixels (for 4K or 8K videos)

Frame Prediction

- Inter Frame Prediction

Frame1

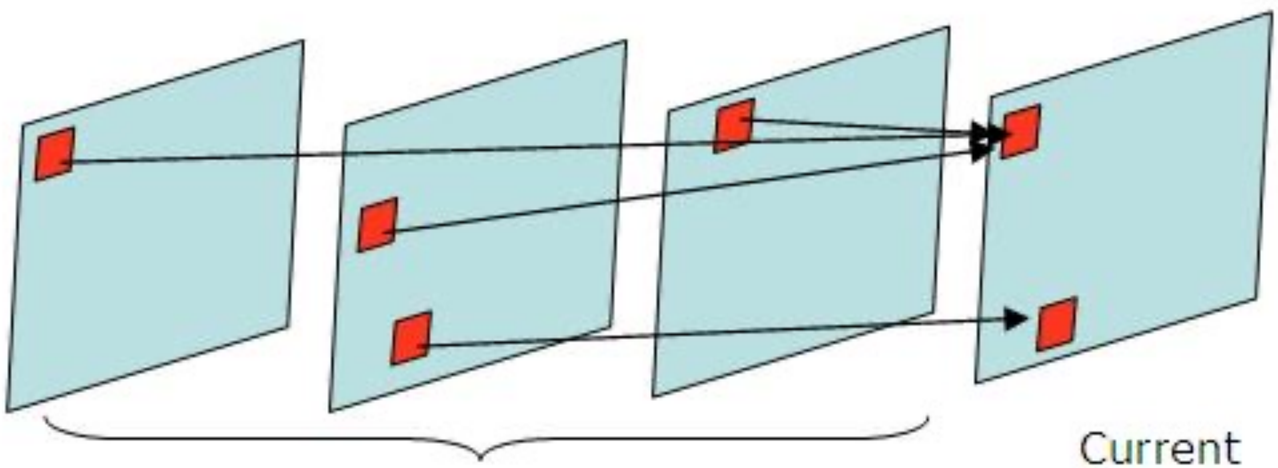
Frame2

Residual – very little ***information***

Source: vcodexer

Inter Frame Prediction

- Instead of directly encoding the raw pixel values for each block, the encoder will try to find a block similar to the one it is encoding on a previously encoded frame, referred to as a reference frame.
- This process is done by a block matching algorithm.



Inter Frame Prediction

- If the encoder succeeds on its search, the block could be encoded by a vector, known as motion vector, which points to the position of the matching block at the reference frame.
- The process of motion vector determination is called motion estimation.

Motion vector visualization

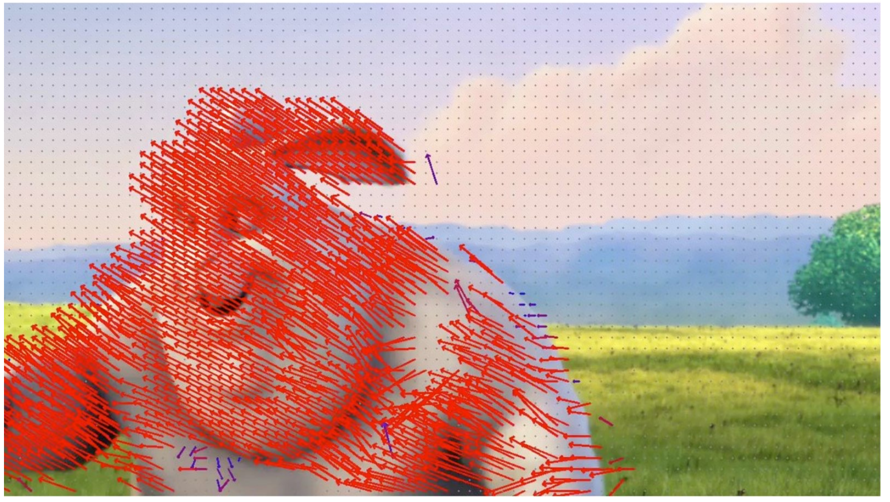
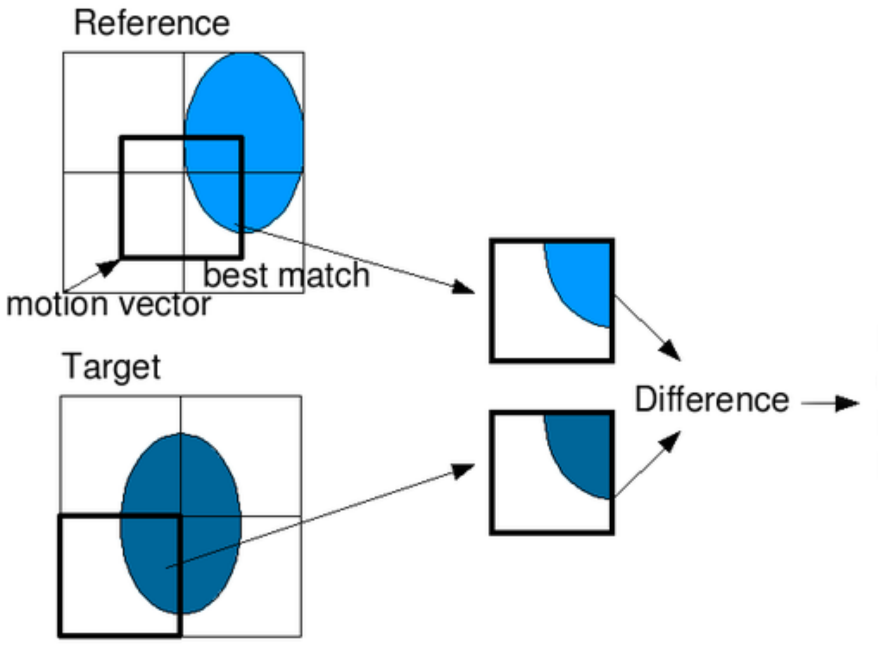


Image credit: Keyi Zhang

Stanford CS348K, Spring 2021

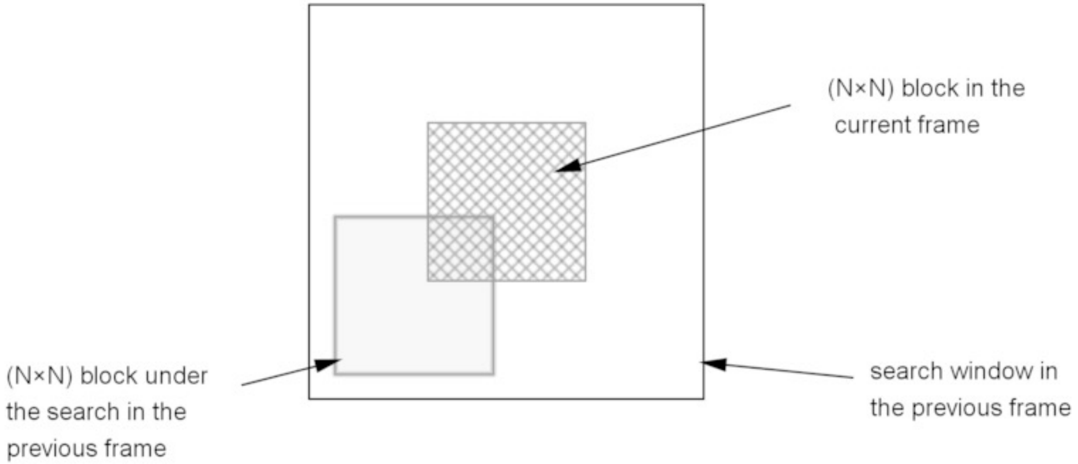
Inter Frame Prediction

- In most cases the encoder will succeed, but the block found is likely not an exact match to the block it is encoding. This is why the encoder will compute the differences between them. Those residual values are known as the prediction error



Inter Frame Prediction

- Block Matching Algorithm



Mean difference or Mean Absolute Difference (MAD) = $\frac{1}{N^2}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}|C_{ij} - R_{ij}|$

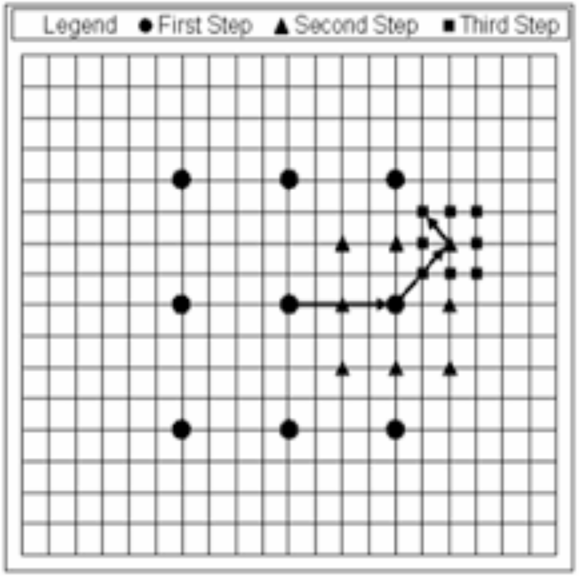
Mean Squared Error (MSE) = $\frac{1}{N^2}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}(C_{ij} - R_{ij})^2$

https://en.wikipedia.org/wiki/Block-matching_algorithm

Inter Frame Prediction

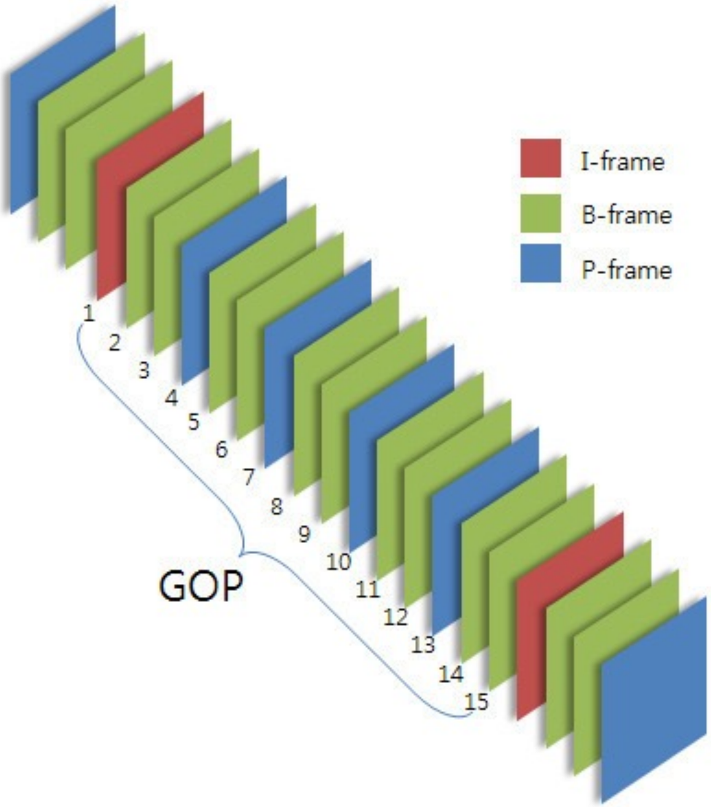
- Types of Block Matching Algorithms
  - Exhaustive search
  - A 3-step search
  - Hexagon or Diamond search

  - Computationally very intensive
  - This must be done for each block of pixels for each frame referencing multiple frames



Inter Frame Prediction

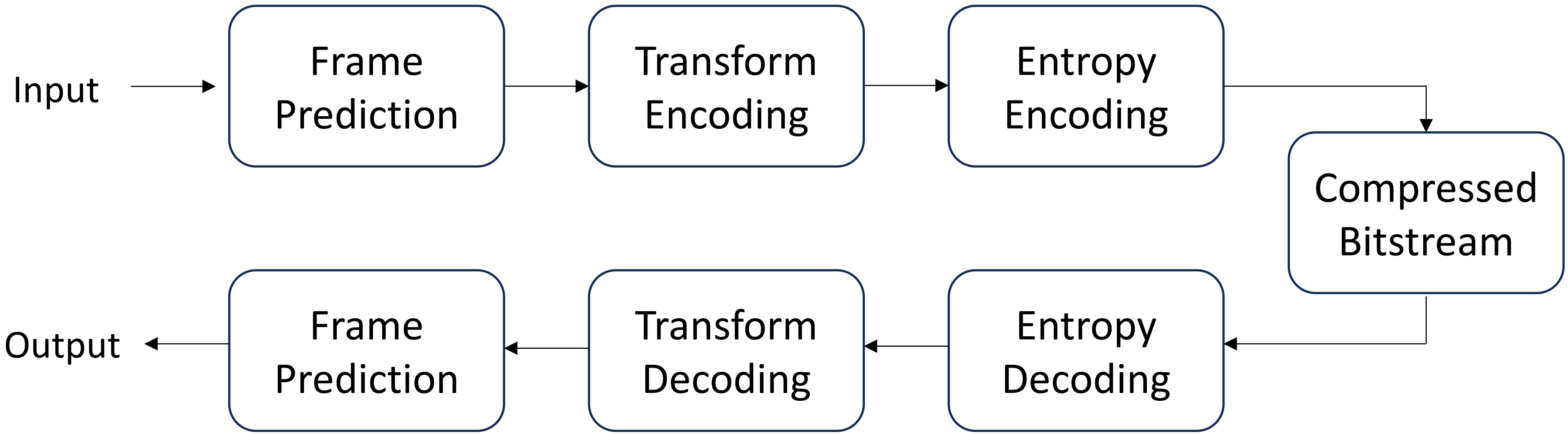
- Three types of frames
  - I – standalone frame, refers itself
  - P – refers to past frames (I or P)
  - B – refers to previous and future frames (P or B)



- Group of pictures (GOP)

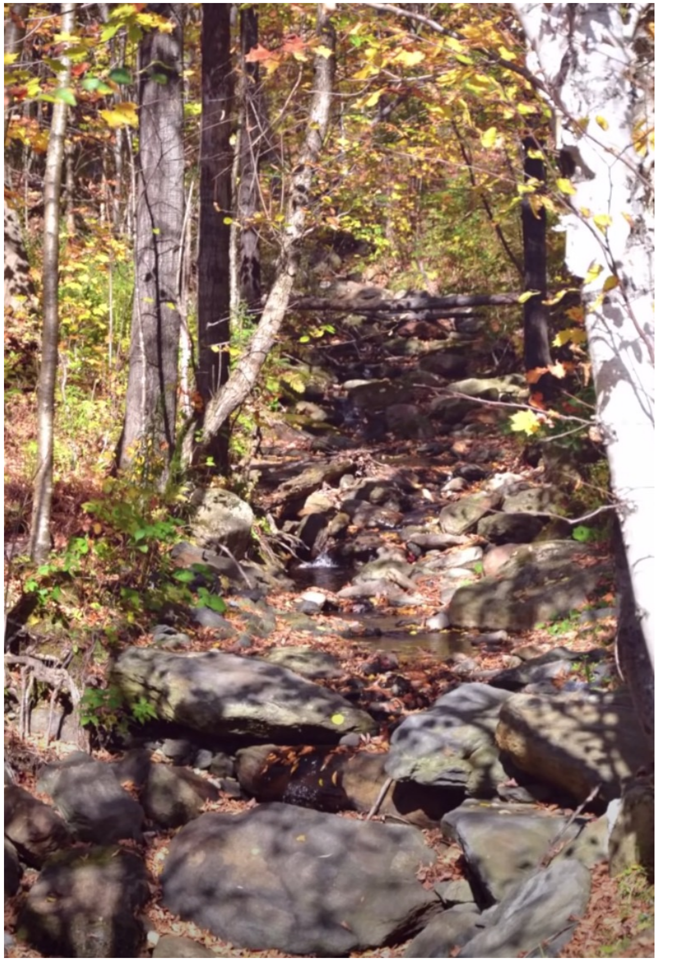
Compression Fundamentals

- Key steps involved in video compression pipeline
  - Color space or Chroma sub-sampling

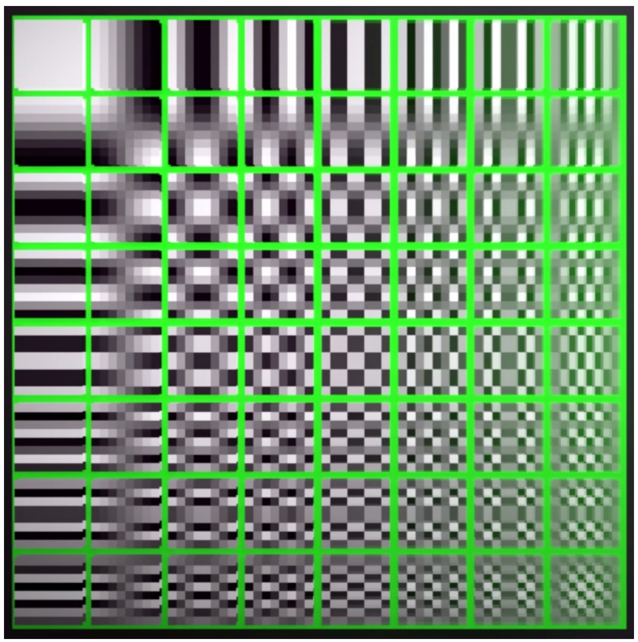


Transform Coding & Quantization

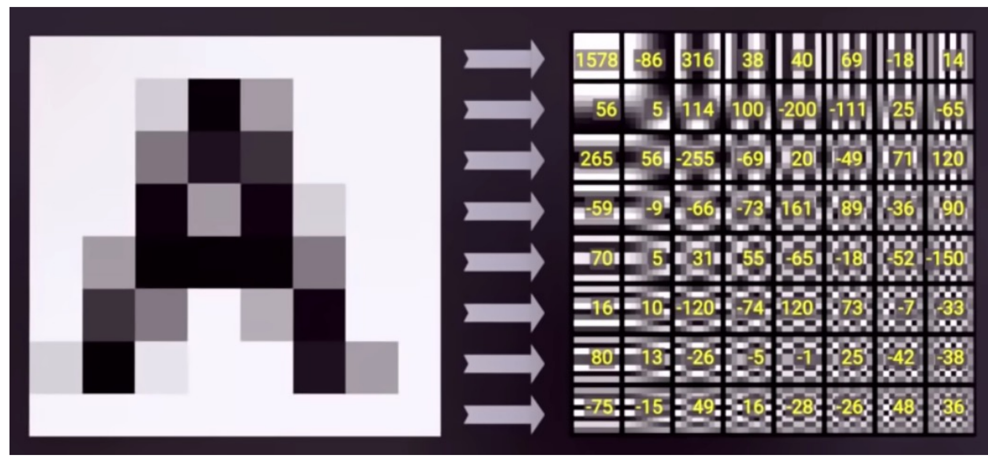
- Transform encoding and quantization
  - Our eyes are bad at perceiving high frequency data
  - Throw away a lot of such data – negligible quality loss



Transform Coding & Quantization



8x8 DCT Transform

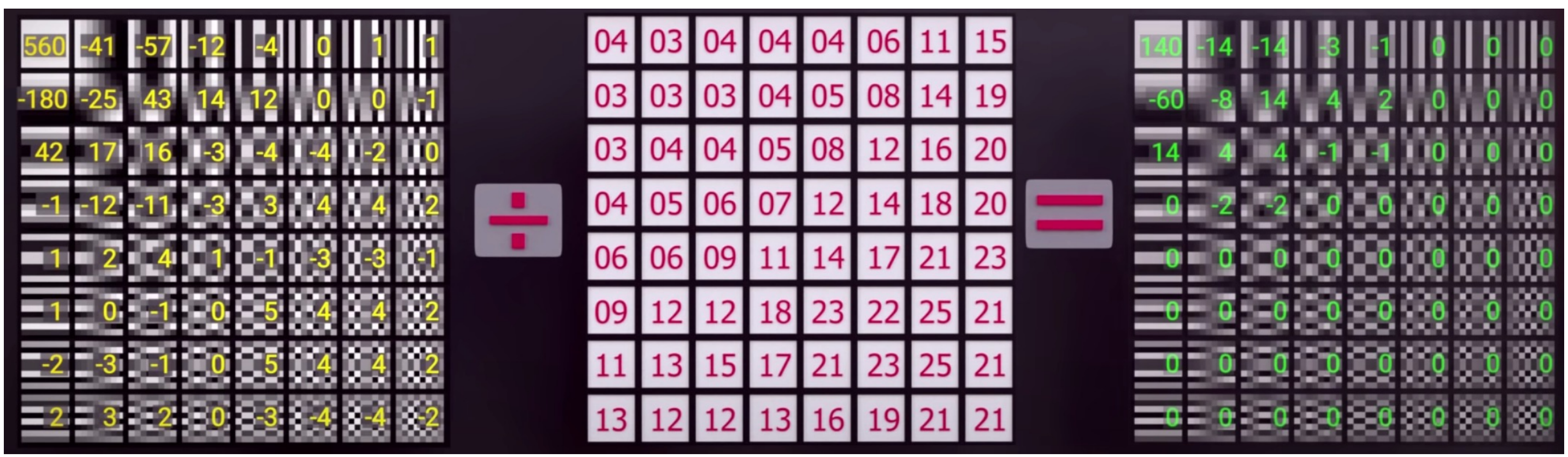


64 constants that represents how much of each base image is used

Transform Coding & Quantization

- Transform encoding and quantization
  - Our eyes are bad at perceiving high frequency data
  - Throw away a lot of such data – negligible quality loss

Lower numbers



DCT'd image

Quantization table

Compressed image

Source: vcodexer

Transform Coding & Quantization

04	04	06	10	21	21	21	21
04	05	06	21	21	21	21	21
06	06	12	21	21	21	21	21
10	14	21	21	21	21	21	21
21	21	21	21	21	21	21	21
21	21	21	21	21	21	21	21
21	21	21	21	21	21	21	21
21	21	21	21	21	21	21	21

Chrominance Quantization Table

Higher numbers generate more 0s

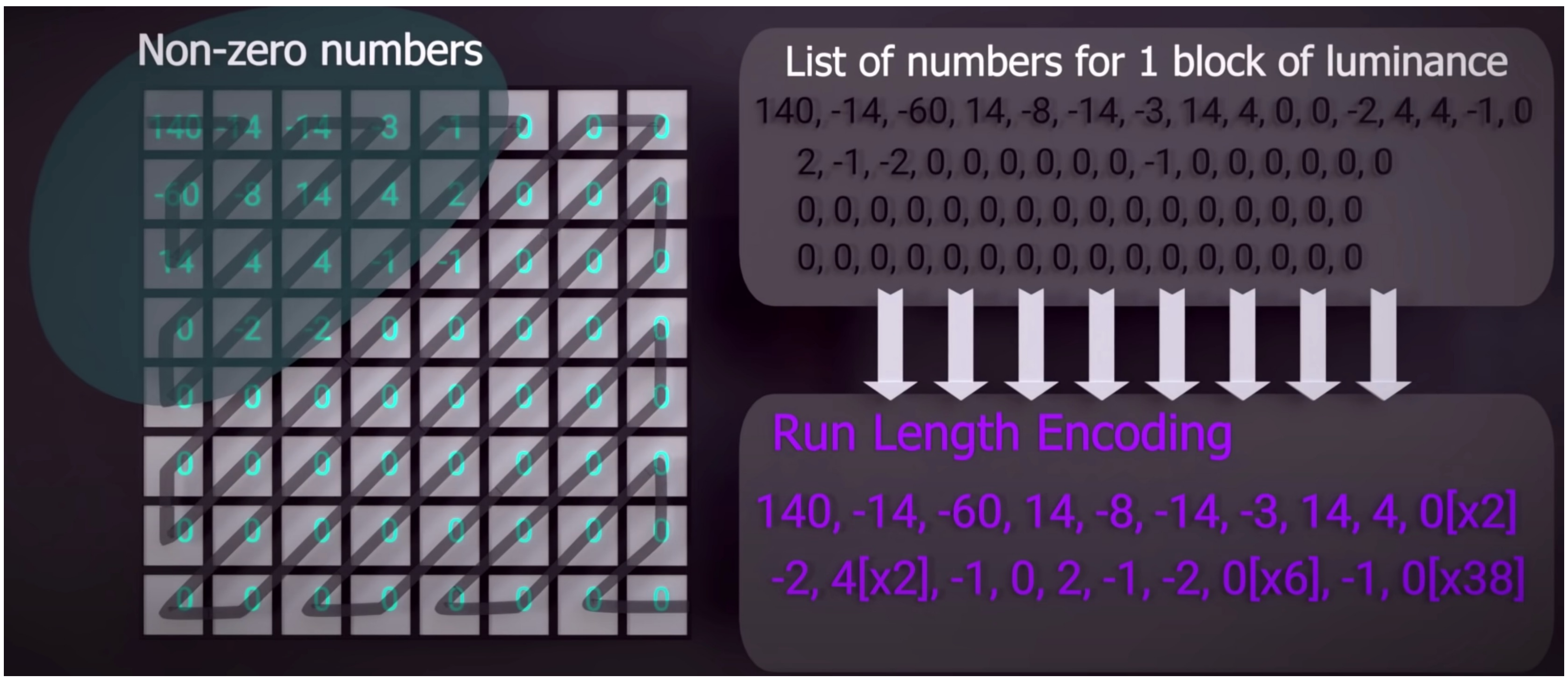
04	03	04	04	04	06	11	15
03	03	03	04	05	08	14	19
03	04	04	05	08	12	16	20
04	05	06	07	12	14	18	20
06	06	09	11	14	17	21	23
09	12	12	18	23	22	25	21
11	13	15	17	21	23	25	21
13	12	12	13	16	19	21	21

Luminance Quantization Table

Lower numbers results in more accuracy

Entropy Coding

- Zigzag Encoding



Entropy Coding

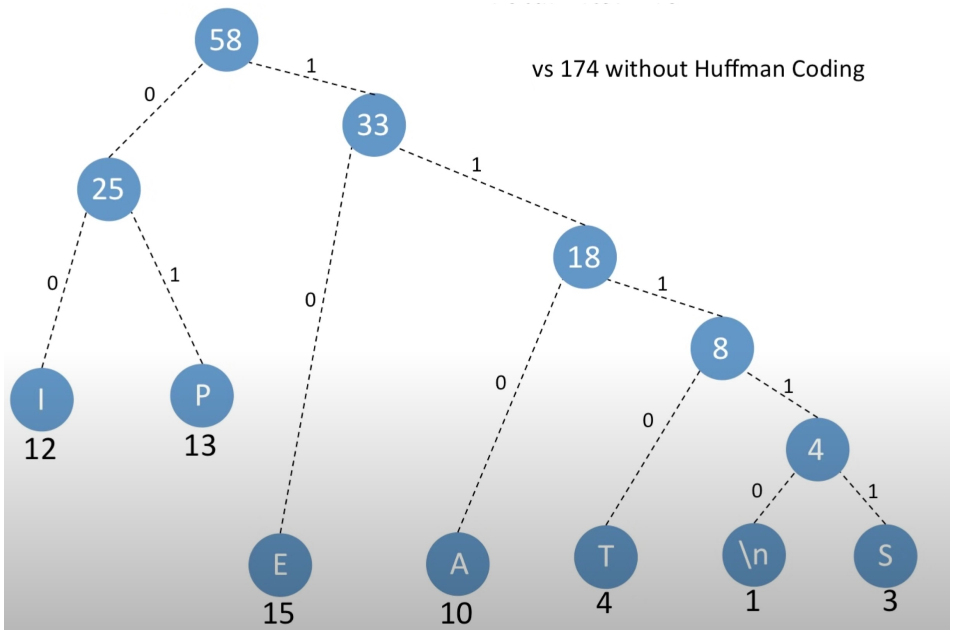
- Huffman coding
  - Based on the lengths of assigned codes on the frequency of data (prefix codes)

Character	Code	Frequency	Total Bits
A	000 Length = 3	10	30 Frequency x Bit Length
E	001	15	45
I	010	12	36
S	011	3	12
T	100	4	12
P	101	13	39
Newline	110	1	3

Total Bits Used: 174

Entropy coding

- Huffman coding



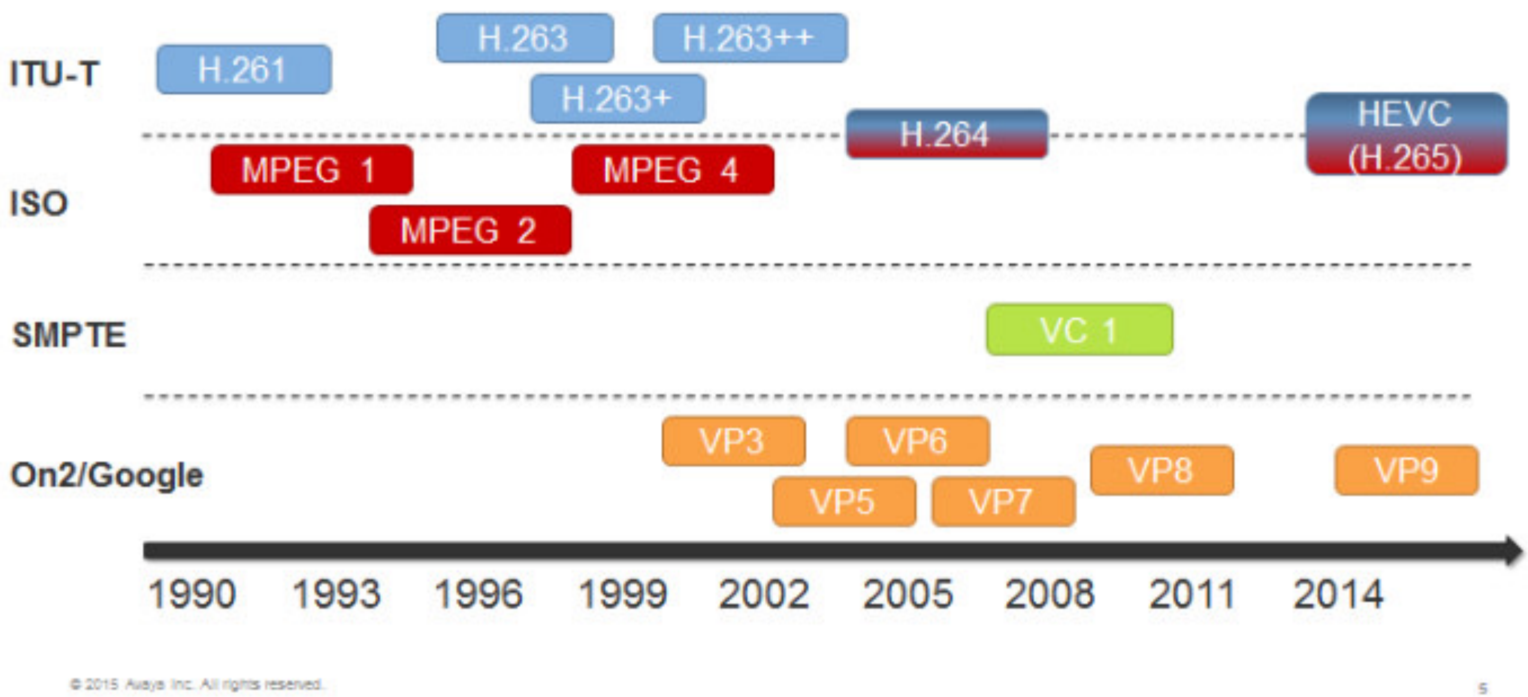
Total Bits: 146

Char	Code	Freq	Total Bits
A	110	10	30
E	10	15	30
I	00	12	24
S	11111	3	15
T	1110	4	16
P	01	13	26
\n	11110	1	5

Compression Artifacts

- 8x8 Blocks

Video Compression History



Popular Video Compression Algorithms

- MPEG Standards
  - MPEG H.26x series, H.266 is the most recent one
  - VP series from Google
  - AV1

Lecture Summary

- Need for Compression
- 2D Compression key steps
  - Chroma sub-sampling
  - Frame prediction
  - Transform coding
  - Entropy coding

Next up: Depth map compression