VLSI DESIGN FOR DIGITAL VIDEO SYSTEMS

Chih-Hung Kuo

Department of Electrical Engineering

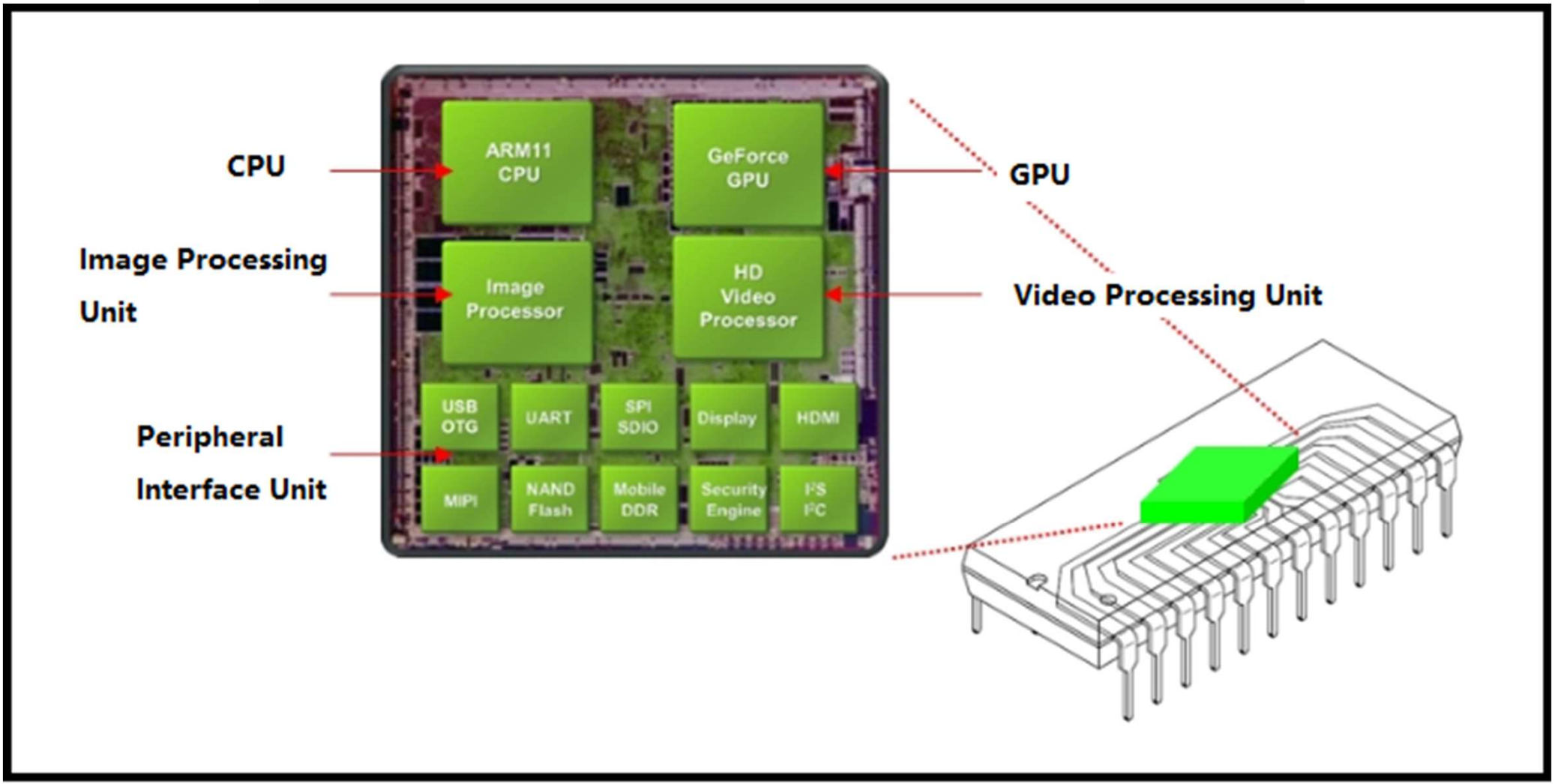
National Cheng Kung University

e-Mail: *chkuo@ee.ncku.edu.tw*

OUTLINE

- VLSI digital signal processing
- Image compression systems
- Video compression systems
- Video processing with deep learning

1. VLSI Digital Signal Processing



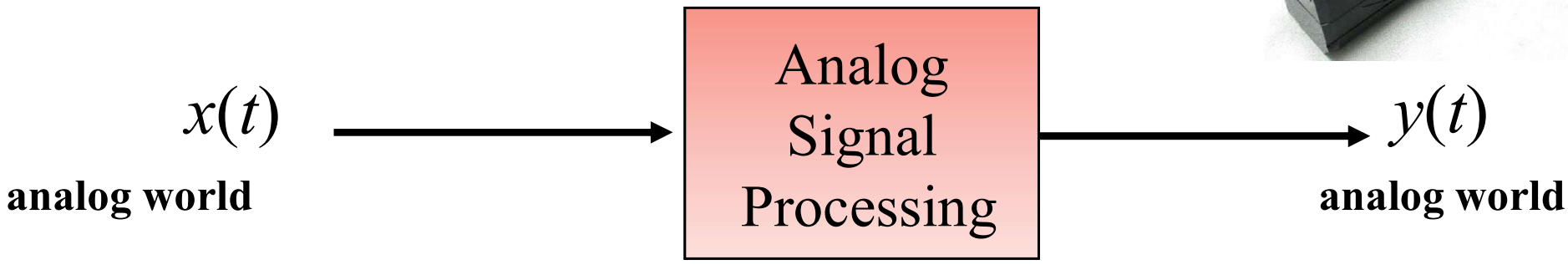
Nvidia Tegra 650 SoC (Source: www.nvidia.com)

SIGNAL PROCESSING SYSTEMS

Input: continuous signal $x(t)$

Output: continuous signal $y(t)$

Processing: analog circuits



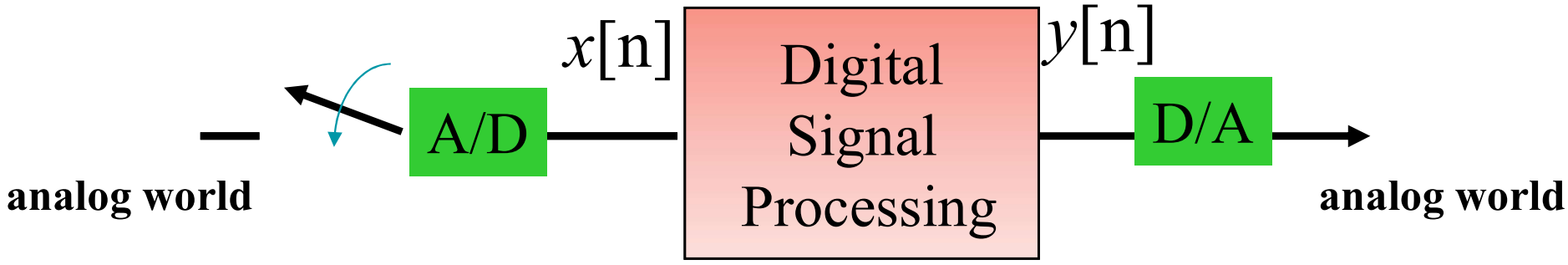
Examples: old TV, radio, telephone, speaker, filter….

SIGNAL PROCESSING SYSTEMS

Input: continuous signal $x(t)$

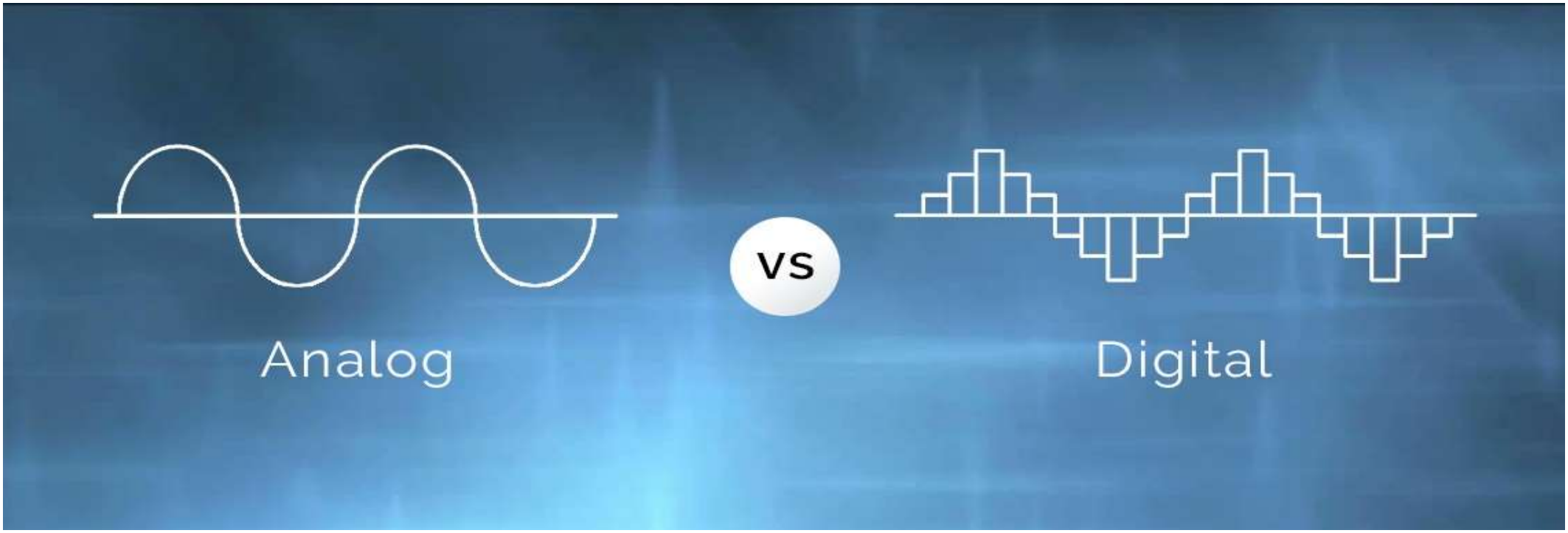
Output: continuous signal $y(t)$

Processing: discrete-time signal $x[n]$, $y[n]$



Examples: smart phone, digital camera, digital video recorder, MP3

TWO WORLDVIEWS



Analog worldview

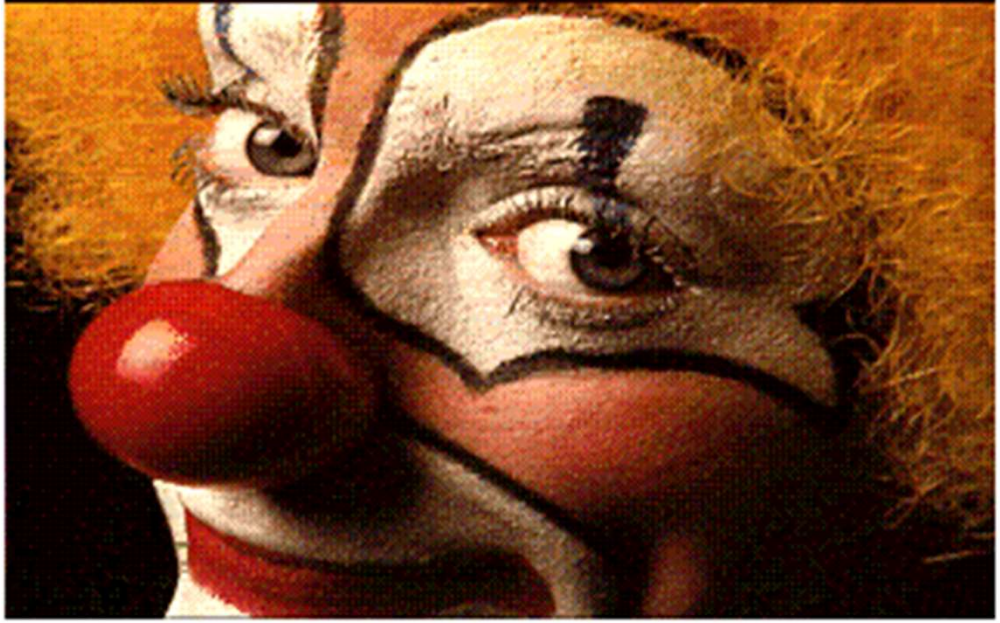
- **calculus**
- **distributions**
- **system theory**
- **electronics**

Digital worldview

- **arithmetic**
- **combinatorics**
- **computing theory**
- **digital signal processing**

2. Image compression

Chih-Hung Kuo



全彩

(無壓縮) 檔案大小：192,054 bytes

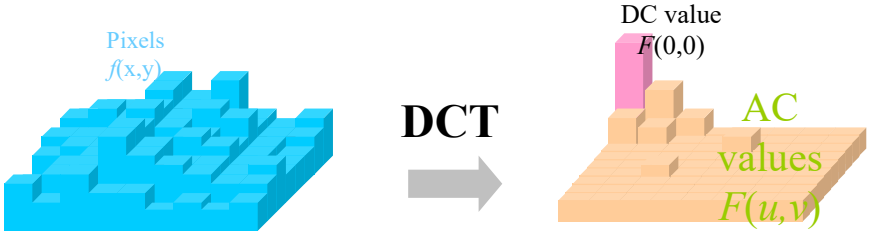
256 色

(無壓縮) 檔案大小：65,078 bytes

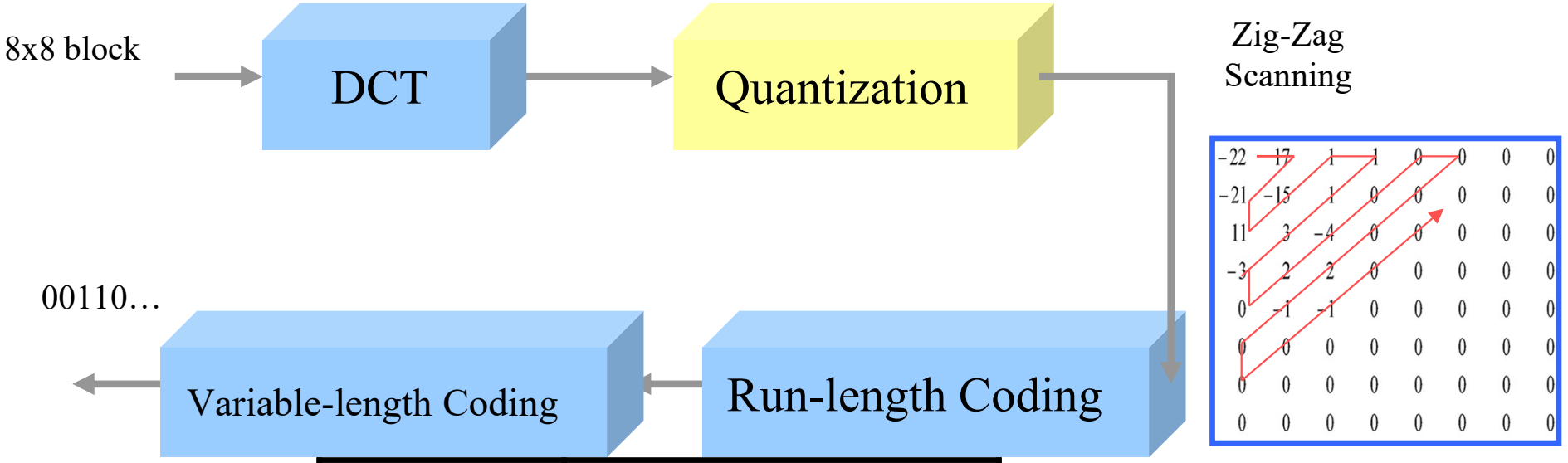
- Full color image

```
( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0)
(221,147, 32) (254,147, 32) (254,179, 32) (254,147, 32) (254,147, 32) (254,147, 32) (254,147, 32) (254,147, 32) (254,254,221) (147,147,147) (254,254,221)
(221,147, 32) (254,147, 74) (254,179, 32) (221,147, 32) (254,179, 32) (254,147, 32) (254,147, 32) (221,147, 32) (254,179,106) ( 74, 74, 74) (254,221,179)
(221,106, 32) (221,147, 32) (254,147, 32) (254,106, 32) (254,147, 32) (221,147, 32) (254,106, 32) (147, 74, 32) (221,106, 32) (254,221,179) (147, 74, 32)
(221,147, 32) (254,147, 32) (221,106, 32) (254,147, 32) (254,147, 32) (221,106, 32) (254,147, 32) (147, 74, 74) ( 74, 32,  0) (106, 74, 32) (254,221,179)
(147, 74, 32) (221,147, 32) (221,147, 32) (254,147, 32) (254,147, 32) (254,147, 32) (254,106, 32) (254,106, 32) (254,147, 32) (254,254,221) (254,221,179)
( 74, 32,  0) (147,106, 32) (221,147, 32) (221,106, 32) (221,147, 32) (221,147, 32) (221,106, 32) (147,106, 74) (254,221,179) (254,221,179) (106, 74, 32)
( 32, 32, 32) (106, 74, 32) (179,147, 32) (221,106, 32) (221,106, 32) (221,106, 32) (179,106, 32) (221,106, 32) (147,106, 74) (254,179,147) (254,179,147)
( 74, 32, 32) (179,106, 32) (179,106, 32) (221,147, 32) (179,106, 32) (221,147, 32) (221,106,  0) (254,147, 32) (254,254,221) (221,179,147) (179,147,106)
( 74, 32,  0) (179,106, 32) (221,147, 32) (221,147, 32) (221,106, 32) (179,106, 32) (221,106, 32) (254,179,147) ( 74, 32, 32) (254,221,179) (254,221,179)
( 32, 32, 32) (106, 32, 32) (179,106, 32) (221,106, 32) (221,147, 32) (179,106, 32) (221,106, 32) (254,179,147) (147, 74, 32) (254,254,221) (254,179,147)
( 32, 32,  0) ( 74, 32, 32) (106, 32, 32) (179,106, 32) (179,106, 32) (147,106, 32) (221,147, 74) (254,179,147) (147, 74, 32) (221,147, 74) (254,179,147)
( 32,  0,  0) ( 32, 32,  0) (106, 74, 32) ( 74, 32, 32) ( 32,  0,  0) (179, 32, 32) (221, 32, 32) (221, 74, 32) (221, 32, 32) (179, 32, 32) (179, 32, 32)
( 32, 32,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) (221, 32, 32) (221, 32, 32) (254,147, 74) (221, 74, 32) (179, 32, 32) (179, 32, 32) (179, 32, 32)
( 32, 32,  0) ( 32,  0,  0) ( 32,  0,  0) (179, 32, 32) (221, 32, 32) (254,179,147) (254,221,147) (254,106, 74) (179, 32, 32) (179, 32, 32) (147, 32, 32)
( 32,  0, 32) ( 32,  0,  0) ( 32,  0,  0) (179, 32, 32) (179, 32, 32) (254,106, 74) (254,106, 74) (179, 32, 32) (179, 32, 32) (147, 32, 32) (106, 32, 32)
( 32,  0, 32) ( 32,  0,  0) ( 32,  0,  0) (179, 32, 32) (179, 32, 32) (179, 32, 32) (179, 32, 32) (147, 32, 32) (147, 32, 32) (106, 74, 32) ( 74, 32,  0)
( 32,  0, 32) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) (147, 32, 32) (147, 32, 32) (147, 32, 32) (106, 32, 32) ( 74, 32,  0) ( 74, 32,  0) ( 32,  0,  0)
( 74,  0, 32) ( 74,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 74, 32,  0) (106, 32, 32) ( 74, 32, 32) ( 74, 32,  0) ( 74, 32,  0) ( 32, 32,  0) ( 32,  0,  0)
( 74,  0, 32) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) (106, 32, 32) ( 32, 32,  0) ( 32, 32,  0) ( 32, 32,  0) ( 32,  0,  0)
( 74,  0, 32) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) (221,179,147) (254,254,221) (221,179,147) (147,106, 74)
( 74,  0, 32) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) (  0,  0,  0) (254,254,221) (179,106, 74) (254,221,179) (221,179,147)
( 74, 32,  0) ( 74,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) (147,147, 74) (147, 32, 32) (147, 32, 32) (106, 32, 32)
( 74, 32,  0) (  0,  0,  0) ( 32,  0,  0) ( 74,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) ( 32,  0,  0) (179,147,106) (179, 32, 32) (147, 32, 32) (179, 32, 32)
( 32, 32,  0) ( 32,  0,  0) ( 32, 32,  0) ( 32, 32,  0) (  0,  0,  0) ( 32,  0,  0) (  0,  0,  0) (254,254,254) (254,254,221) (179, 32, 32) (179, 74, 32)
```

JPEG - BASIC TRANSFORM CODING



$F'(u,v)=\text{round}[F(u,v)/Q(u,v)]$



(run, amplitude)	Variable-length codeword
(0,12)	0000 0000 1101 00
(0,6)	0010 0001 0
(1,4)	0000 0011 000
(0,3)	0010 10
EOB	10

Run 0 0 0 0 0 0 0 0 0 0 1 0 5 0 3 EOB

Amplitude -22 17 -21 11 –15 1 1 1 3 -3 2 -4 2 -1 -1 EOB

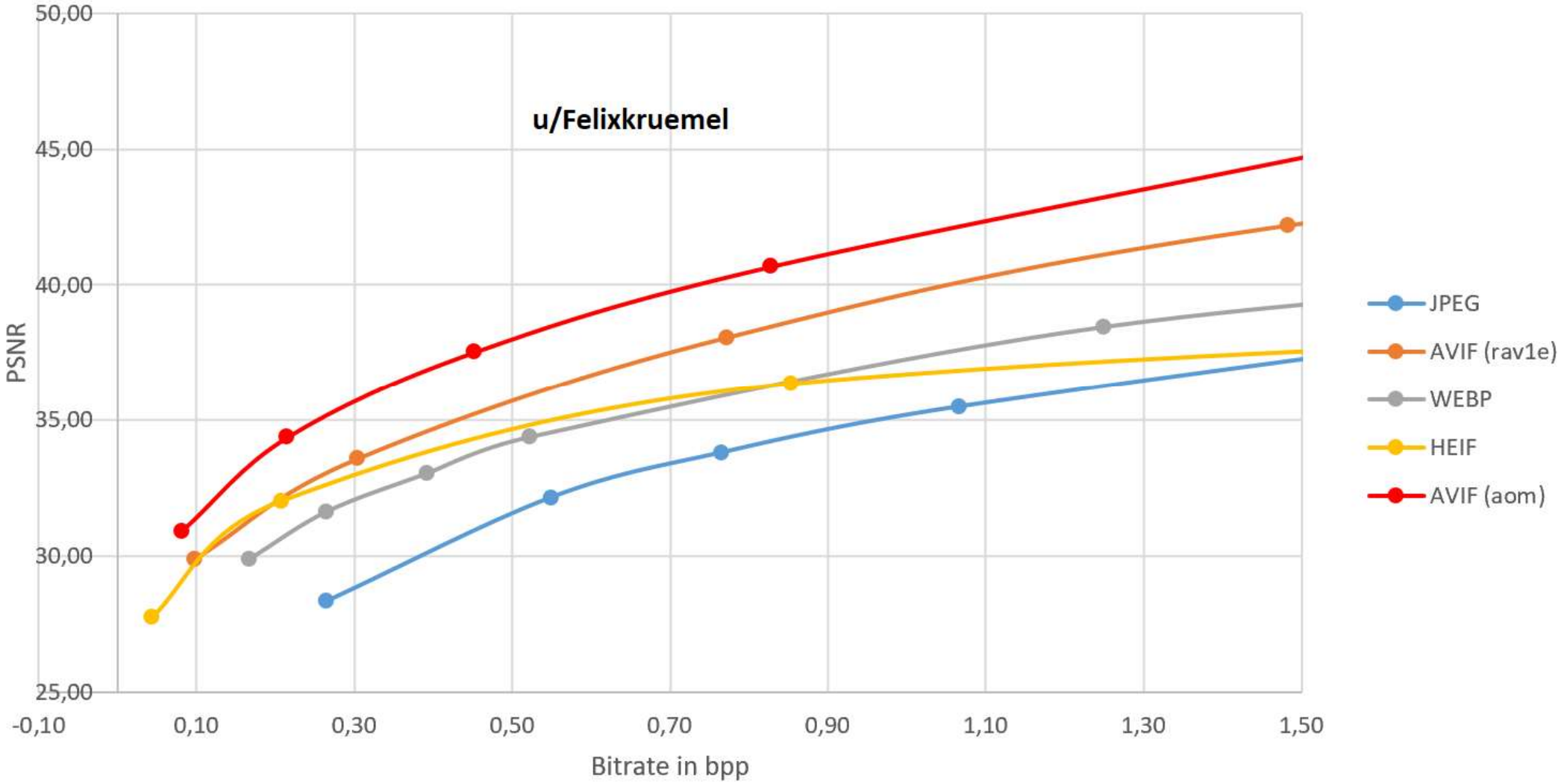
IMAGE CODING STANDARDS

- **JPEG**
  - Joint Photographic Experts Group
- **JPEG 2000**
  - Progressive coding based on wavelet transform
- **WebP**
  - Google, available in all browsers
- **HEIC**
  - MPEG group, based on HEVC
- **AVIF**
  - Alliance for Open Media (AOM), Chrome and Firefox
- **JPEG XL**
  - JPEG group, the next-generation codec
- **WebP2**
  - Google, an experimental successor to WebP

QUALITY COMPARISON

https://jakearchibald.com/2020/avif-has-landed/demos/compare/

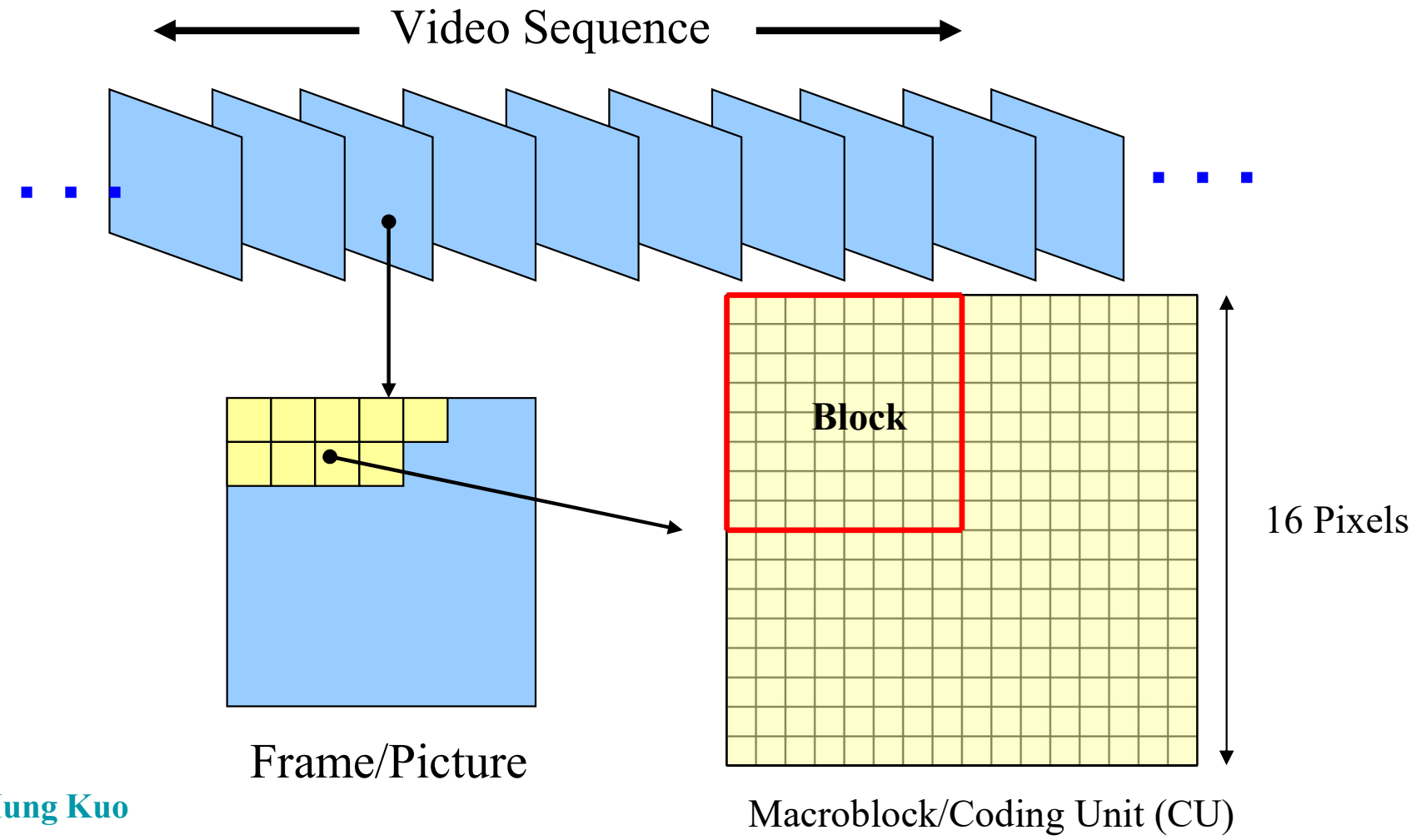
QUALITY COMPARISON



3. Video Compression

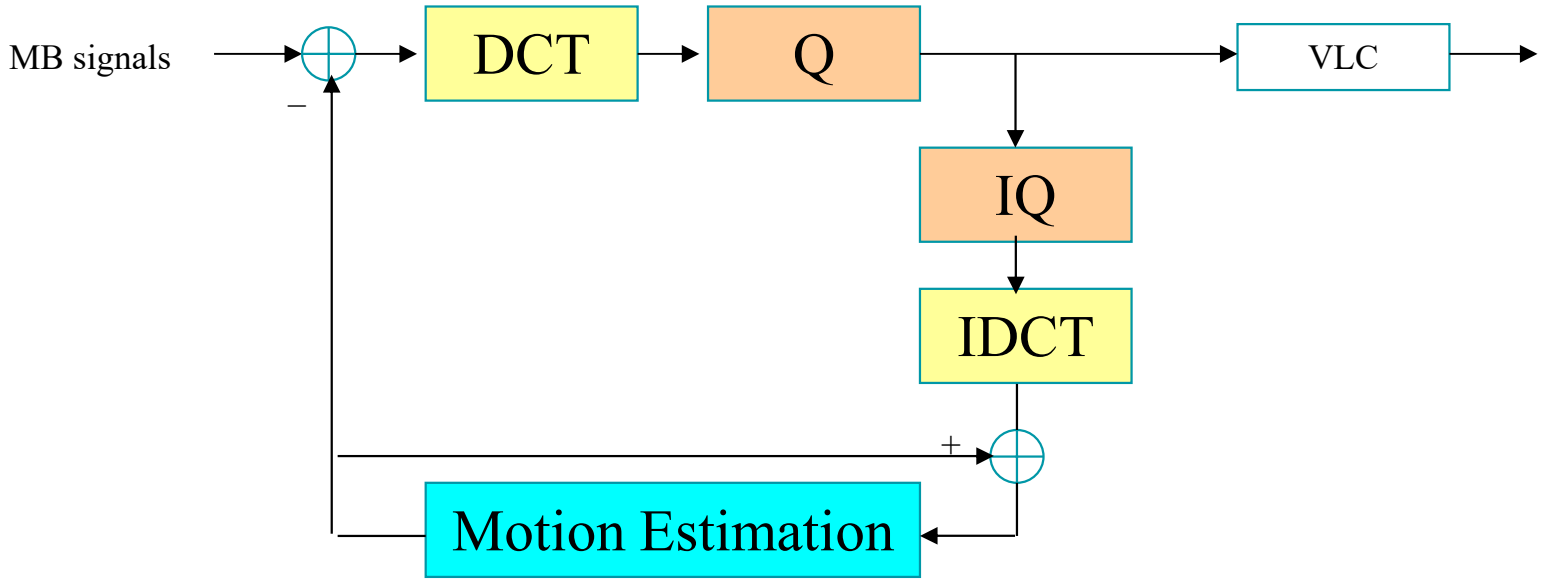
Chih-Hung Kuo

VIDEO STREAM DATA



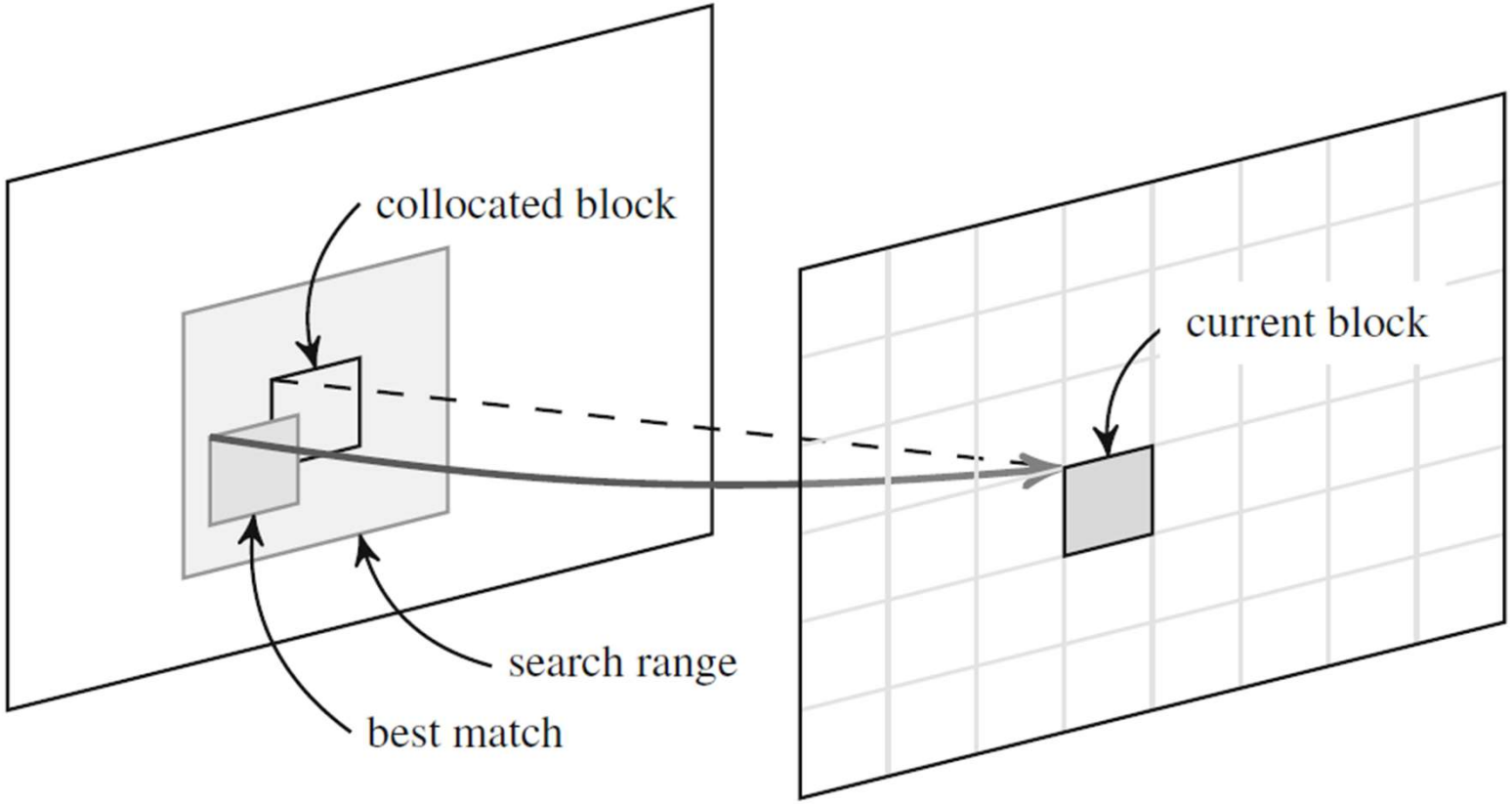
COMMON CODING COMPONENTS

- Each image frame is divided into macro-blocks (MB)
- Video compression: reduce redundancies in image sequences
  - Spatial domain: Digital Cosine Transform(DCT), intra prediction
  - Temporal domain: motion estimation
  - Quantization and entropy coding



MOTION ESTIMATION AND COMPENSATION

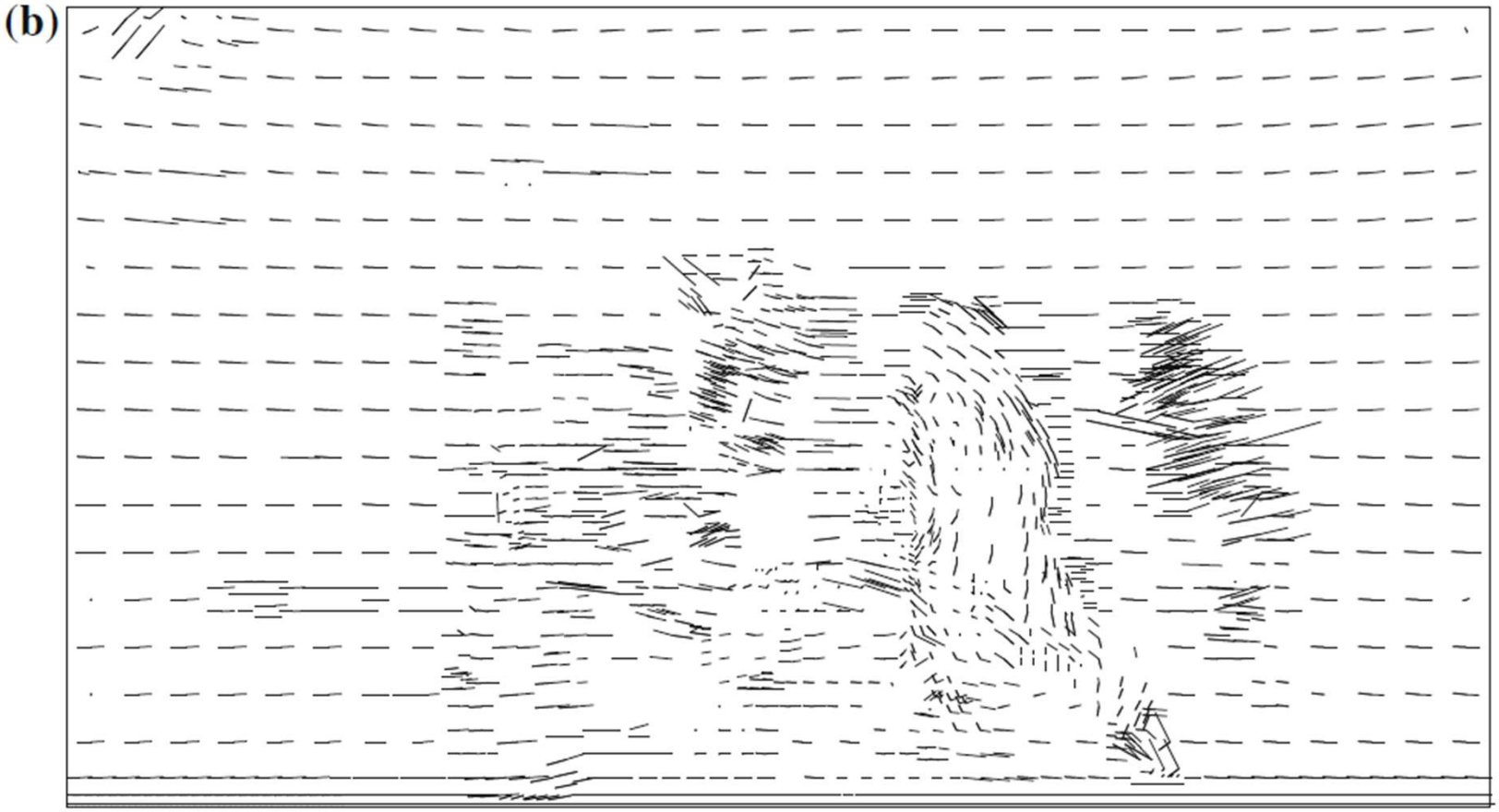
- Motion estimation in a search range around the collocated block in the reference picture



Reference : “High Efficiency Video Coding” by Mathias Wien

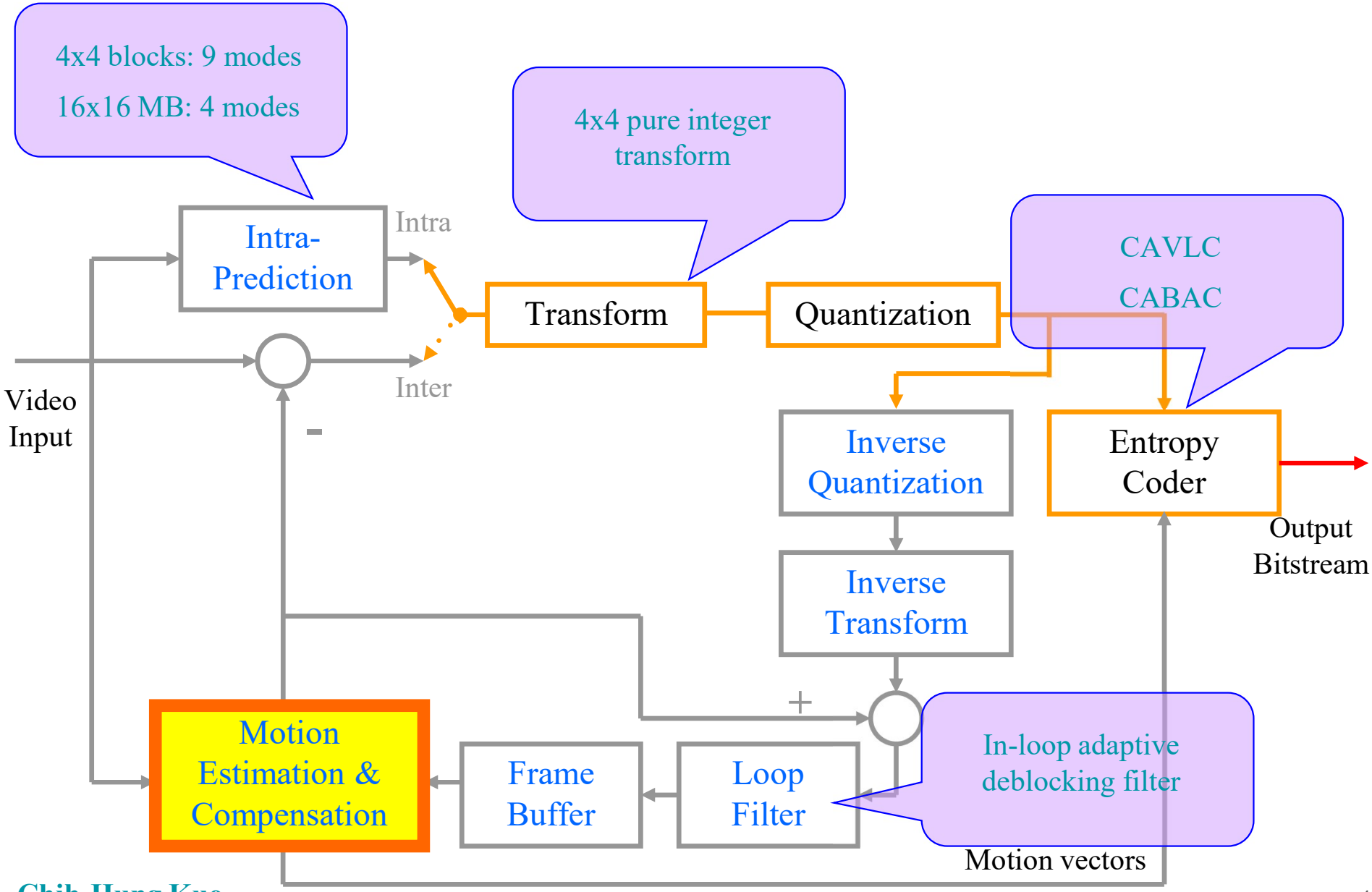
MOTION ESTIMATION AND COMPENSATION

- Motion for a complete picture



Reference : “High Efficiency Video Coding” by Mathias Wien

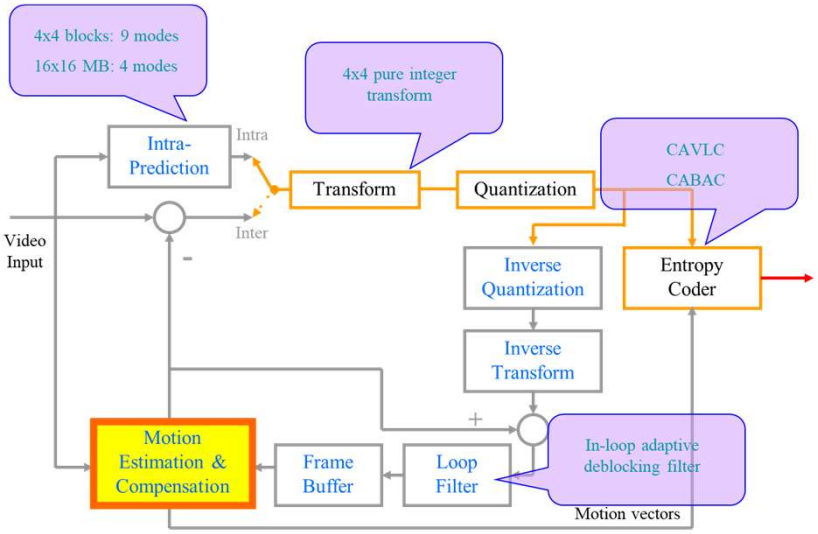
H.264/AVC ENCODER



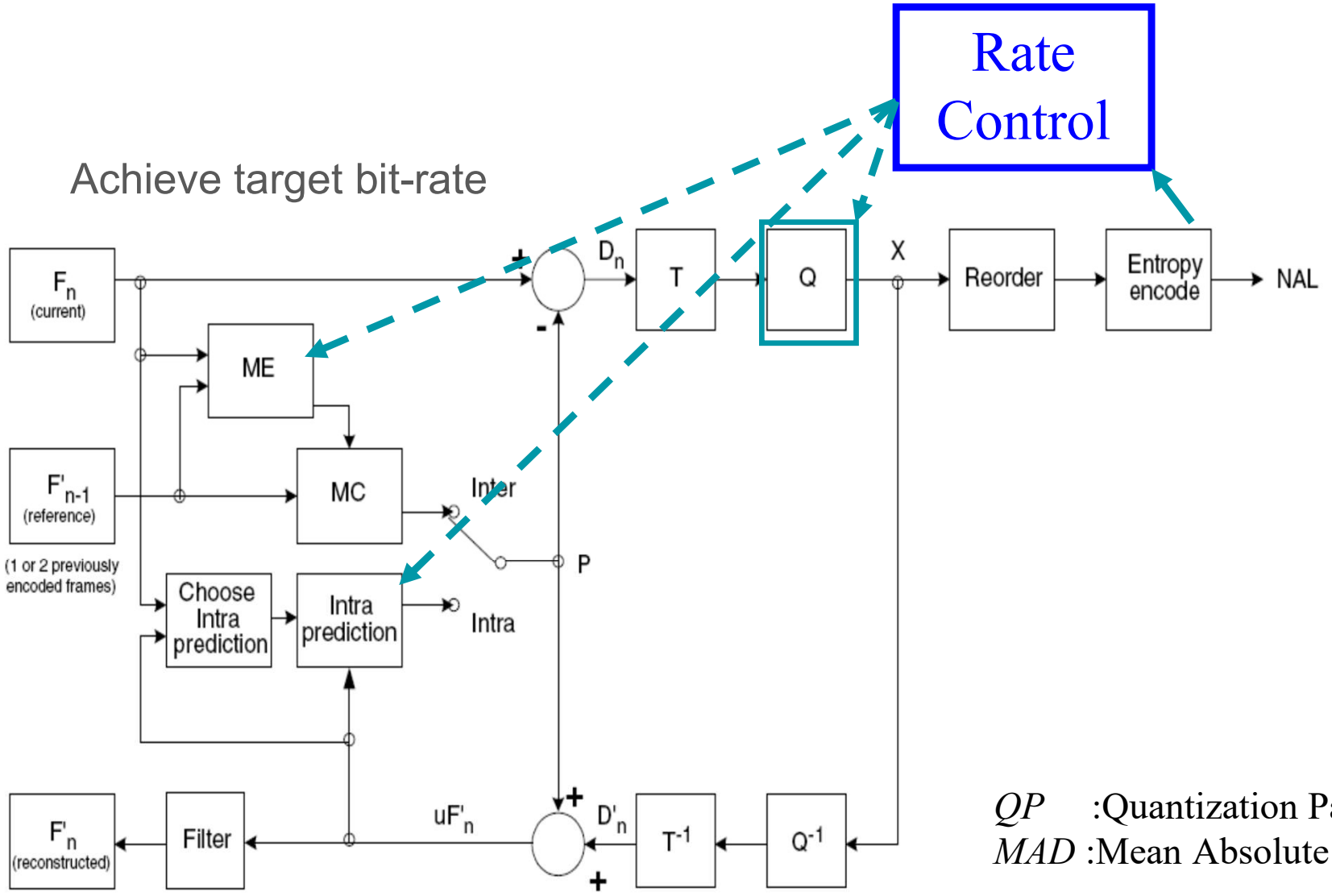
https://youtu.be/K7YSzd6JzLI

DESIGN ISSUES

- Efficient component design
  - Entropy coding, DCT, motion estimation, …
- High complexity
  - “The tyranny of numbers”
  - Hardware/software co-design
  - **System level design**



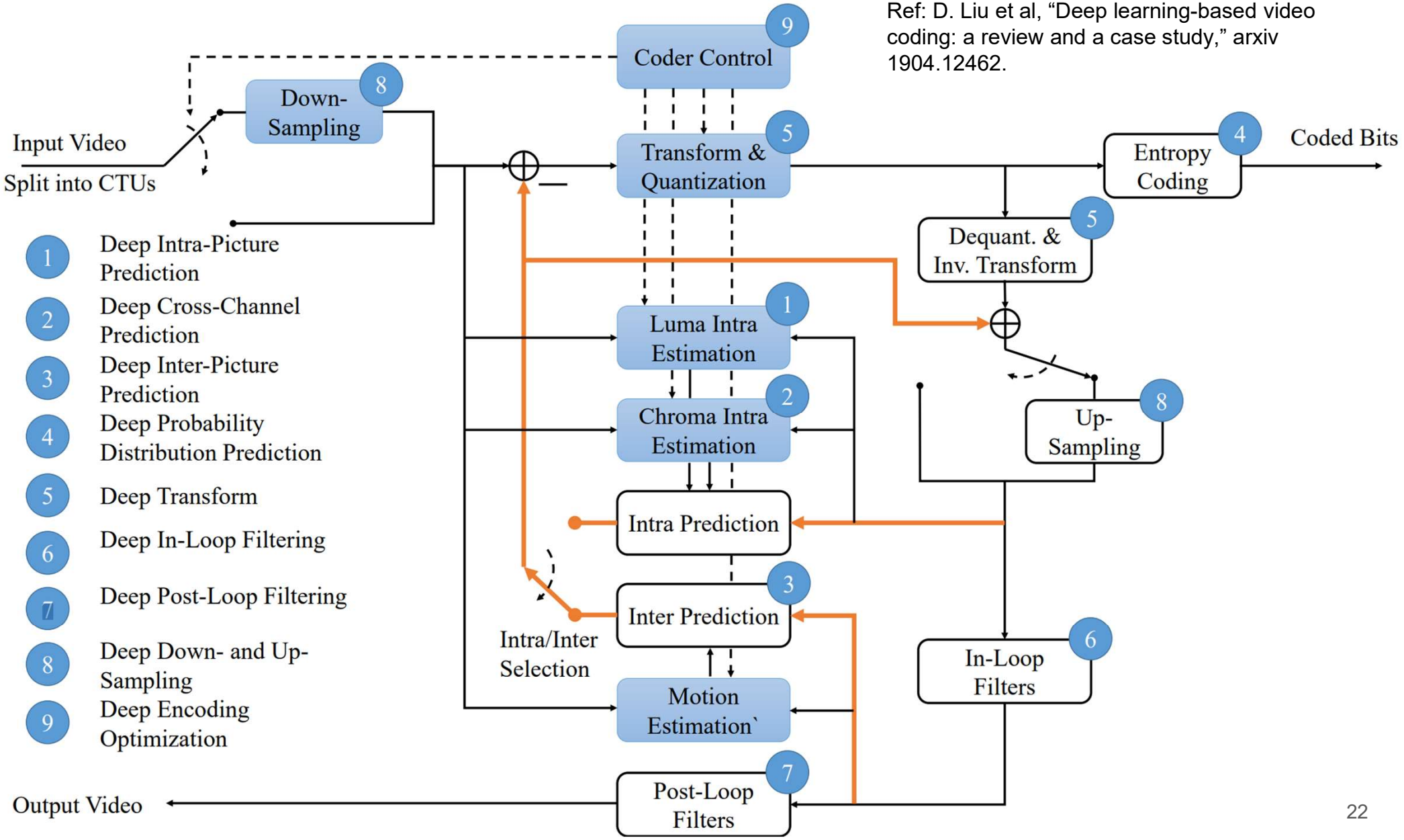
Example: rate control for video coding



QP :Quantization Parameter
MAD :Mean Absolute Difference

AI-assisted video coding

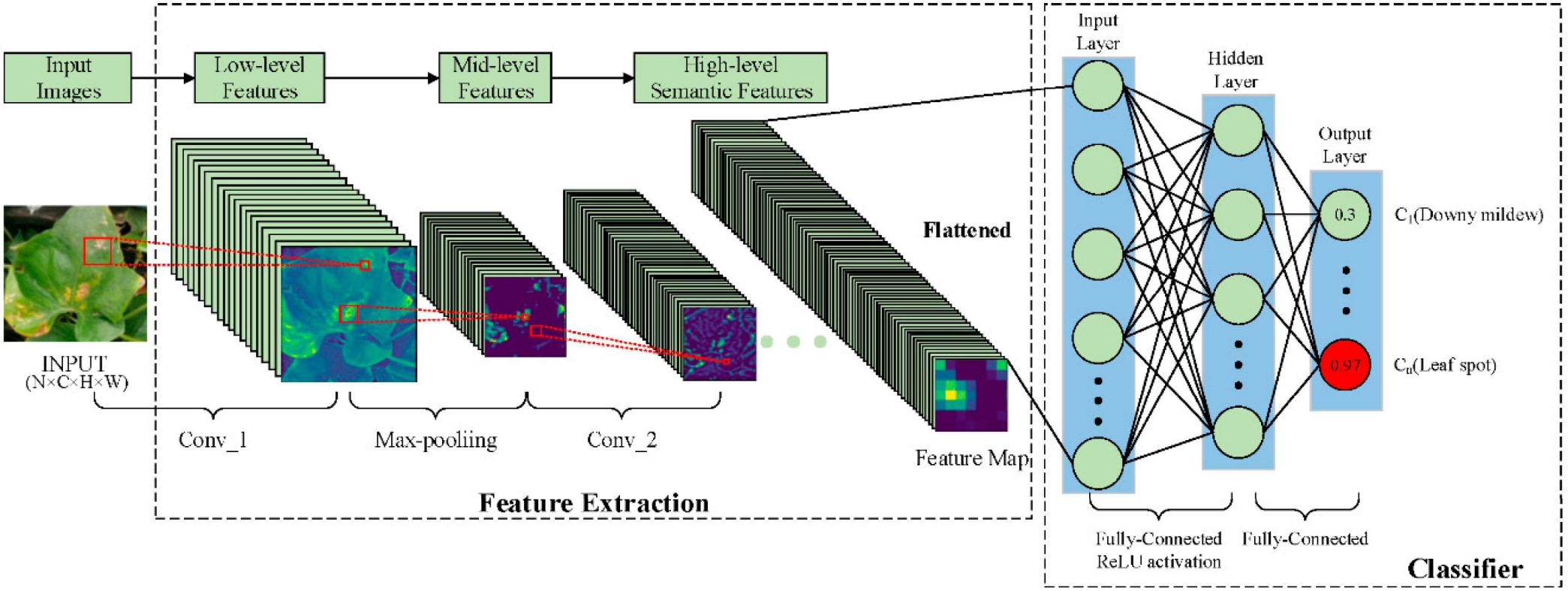
Ref: D. Liu et al, “Deep learning-based video coding: a review and a case study,” arxiv 1904.12462.



4. Deep Learning

Chih-Hung Kuo

Deep Learning Model

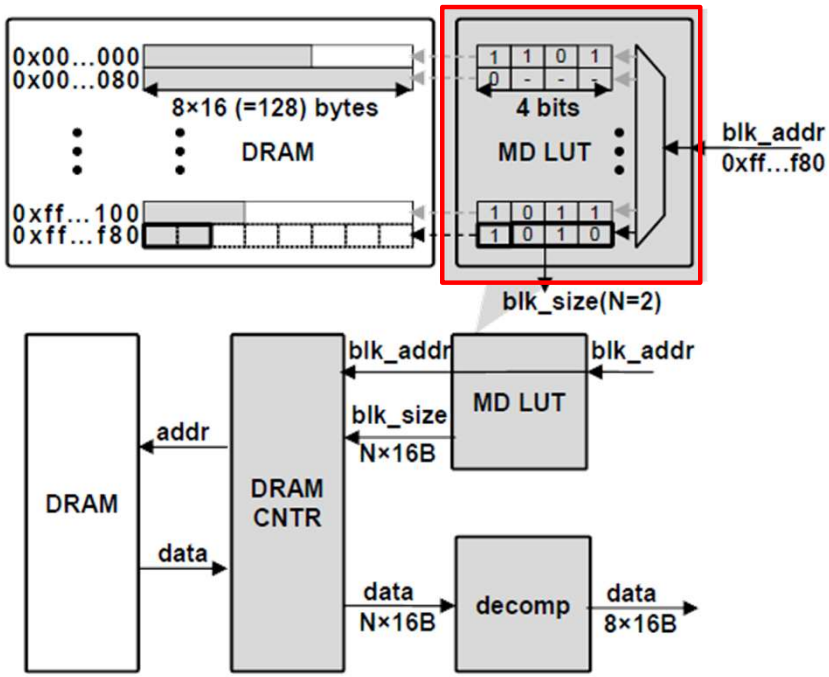


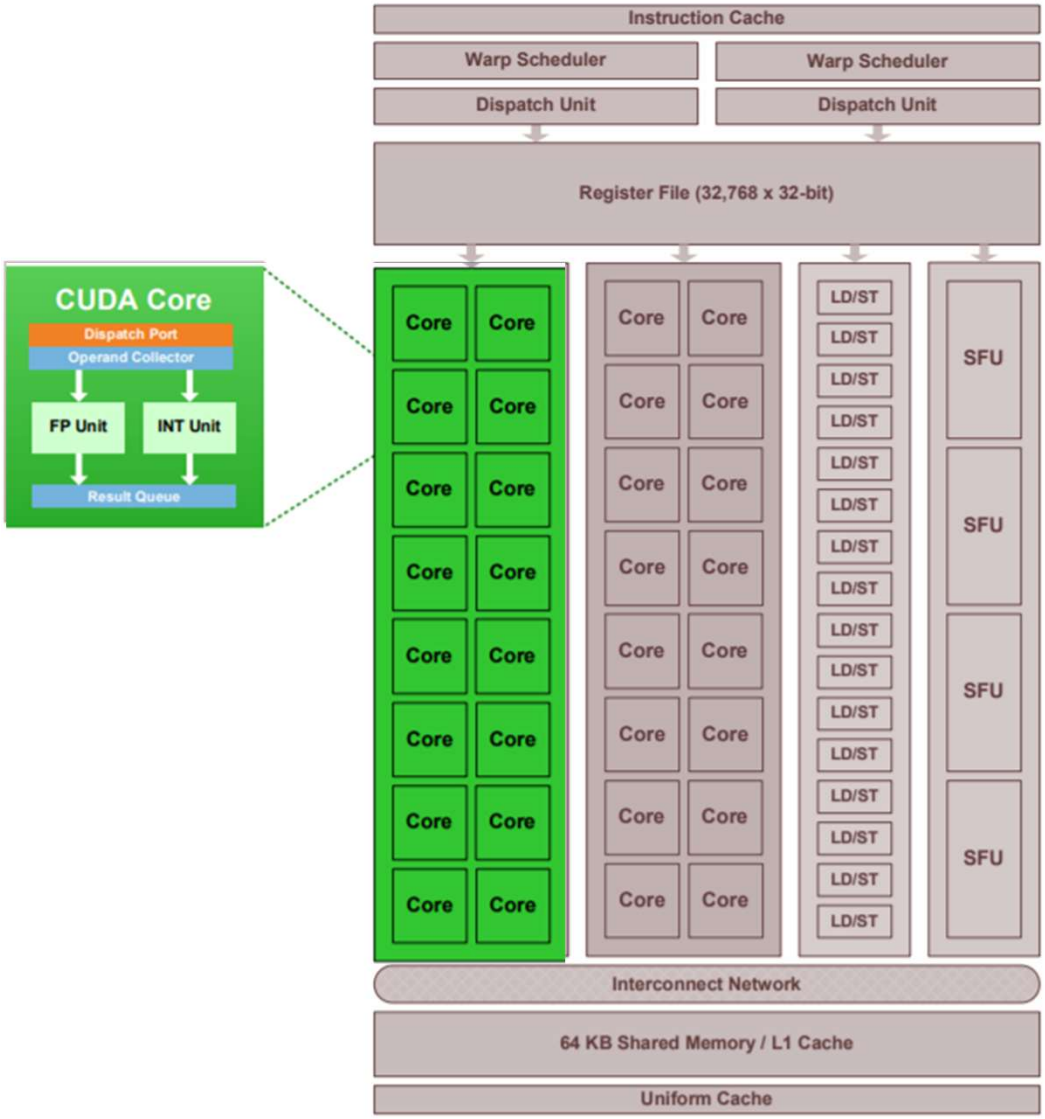
Source: https://www.mdpi.com/2077-0472/11/8/707

Circuits for Deep Learning

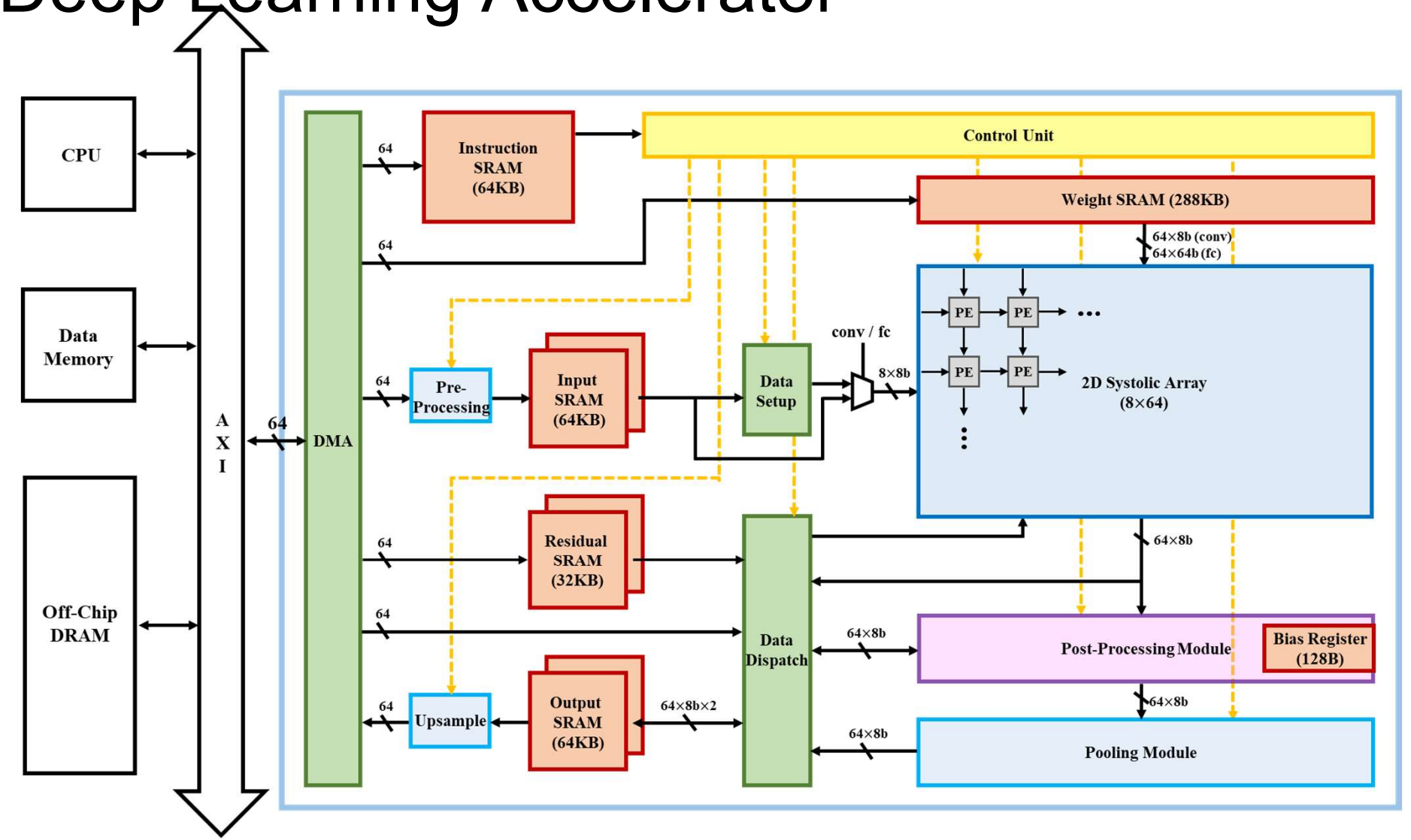
- GPU: Graphics processing unit

Memory I/O compression





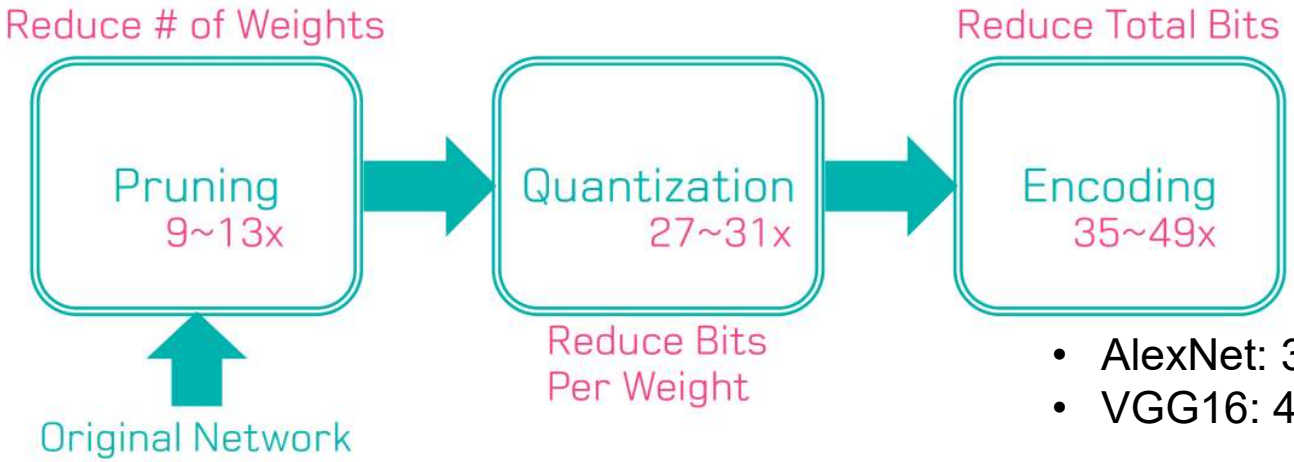
Deep Learning Accelerator



Systolic Array:

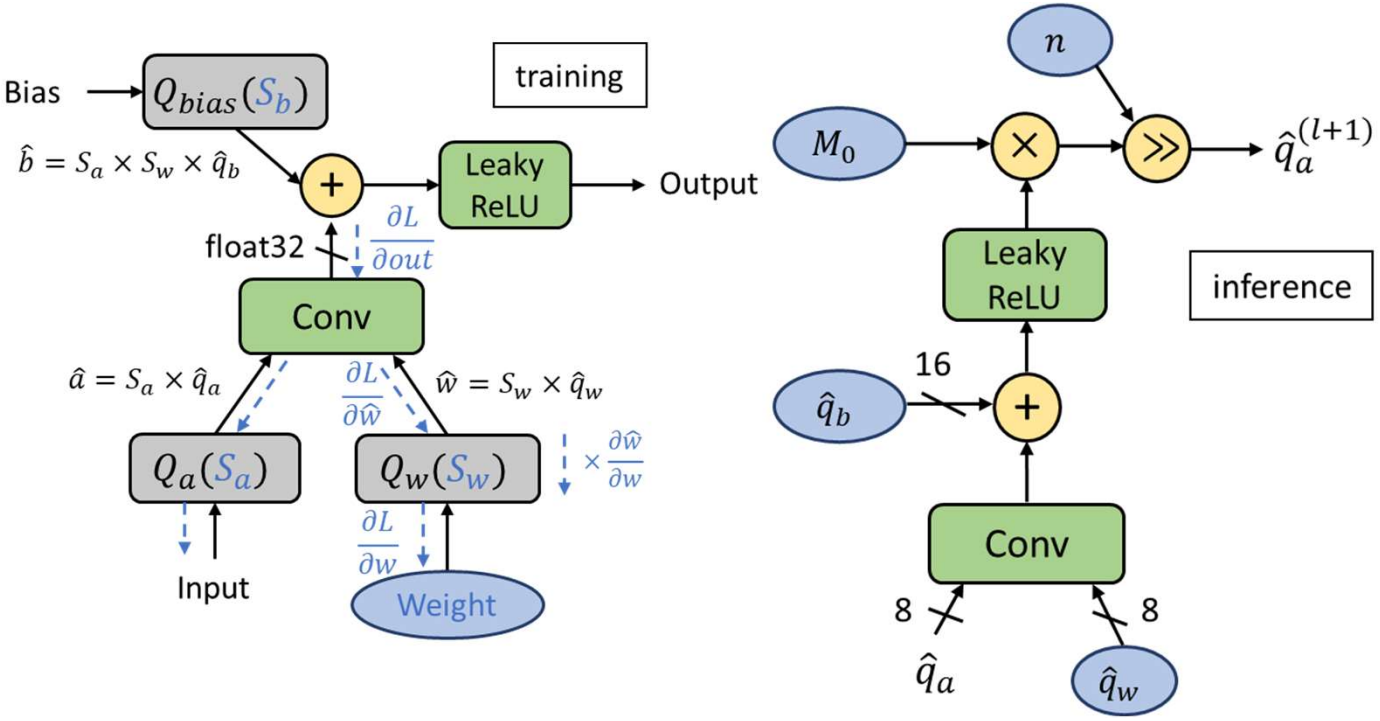
- Simple and regular design
- Concurrency and communication
- Balancing computation with I/O

Neural network compression



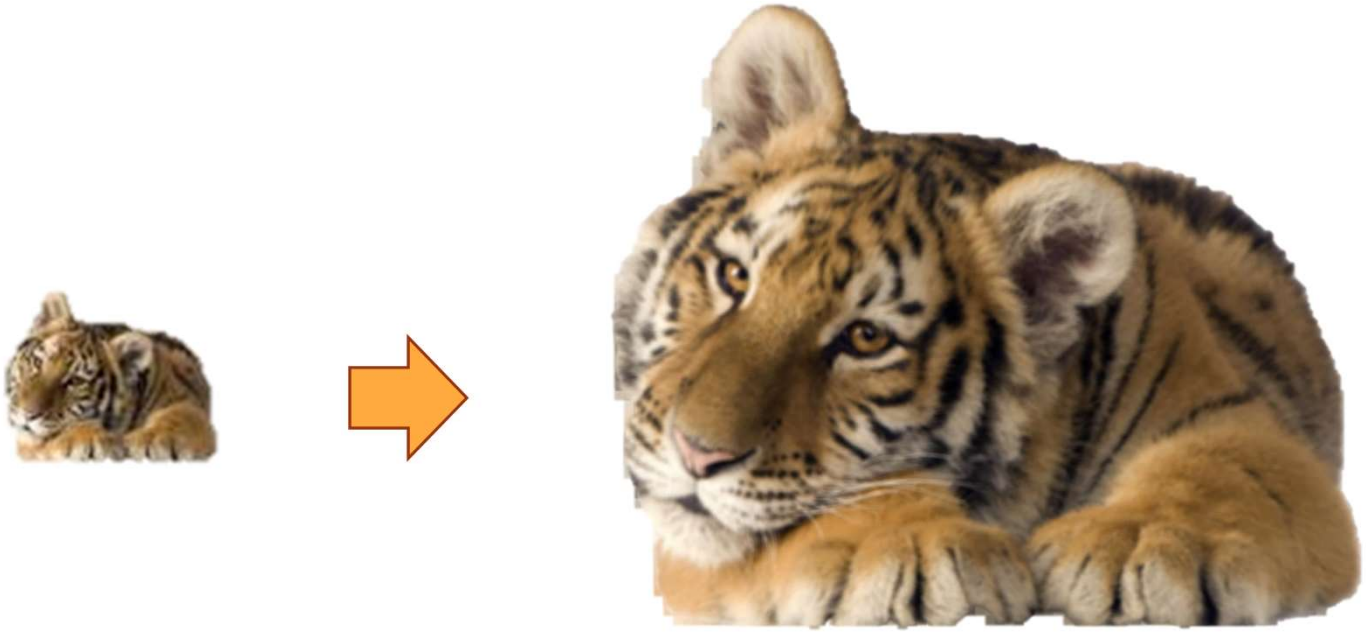
- AlexNet: 35×, 240MB => 6.9MB
- VGG16: 49× 552MB => 11.3MB

Quantization-aware training



Application: Image Super-Resolution

- Input low-resolution image to get a larger and clear one



Super-Resolution Performance

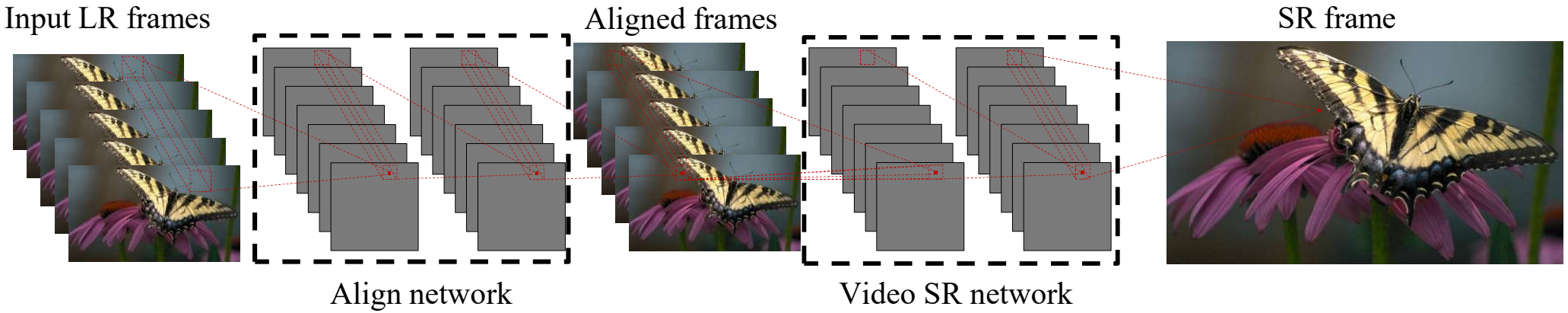
- Set14 “Baboon” x4

Bicubic 22.44 / 0.4510

Super-Resolution 24.86 / 0.6953

Video Super-Resolution

1. Input multiple input frames
2. Perform alignment for each input frame
3. Concatenation of all Aligned frames and input to Video SR network



Video Super-Resolution Demo - city

SR Magnified

Bicubic interpolation

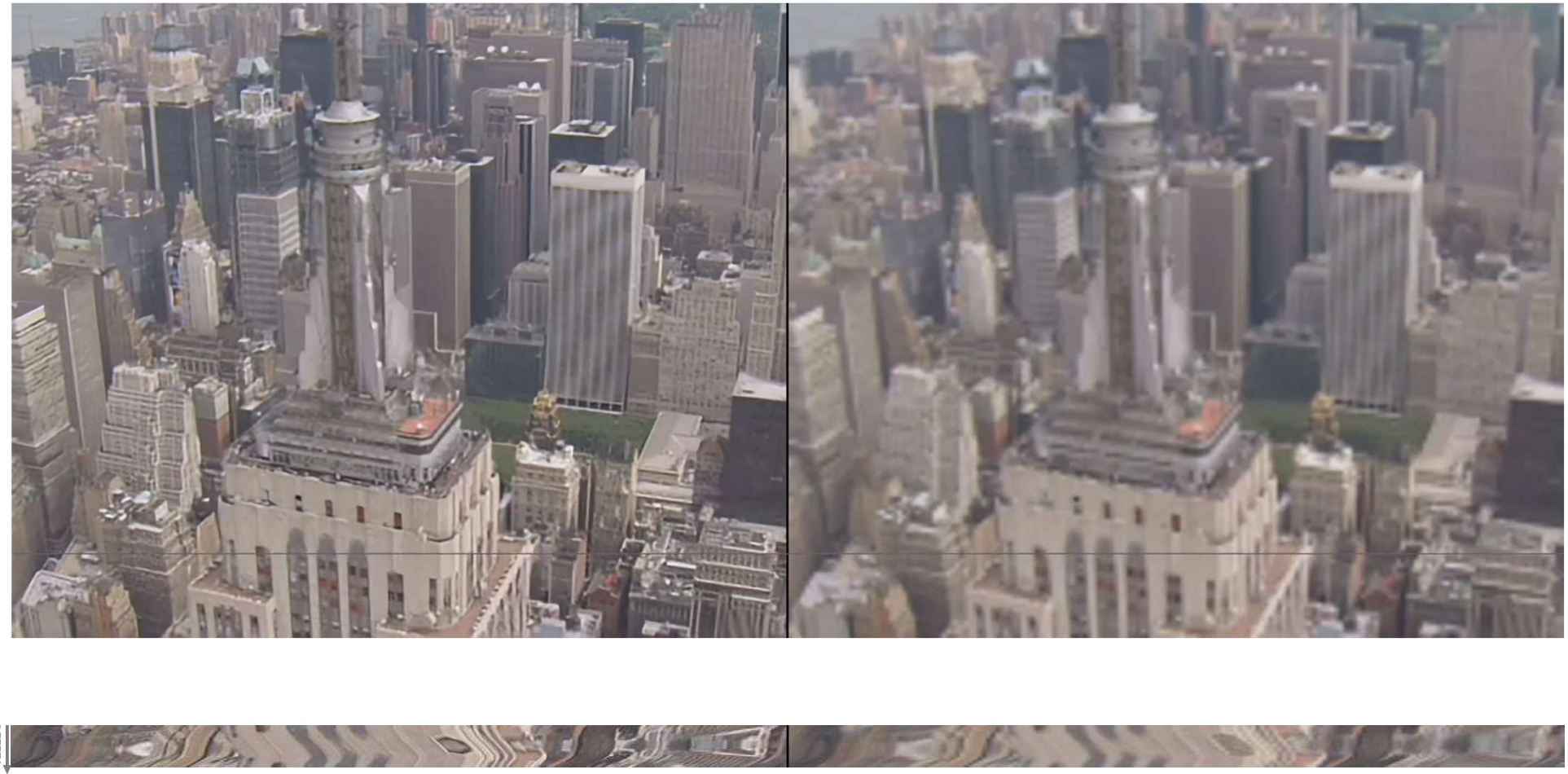
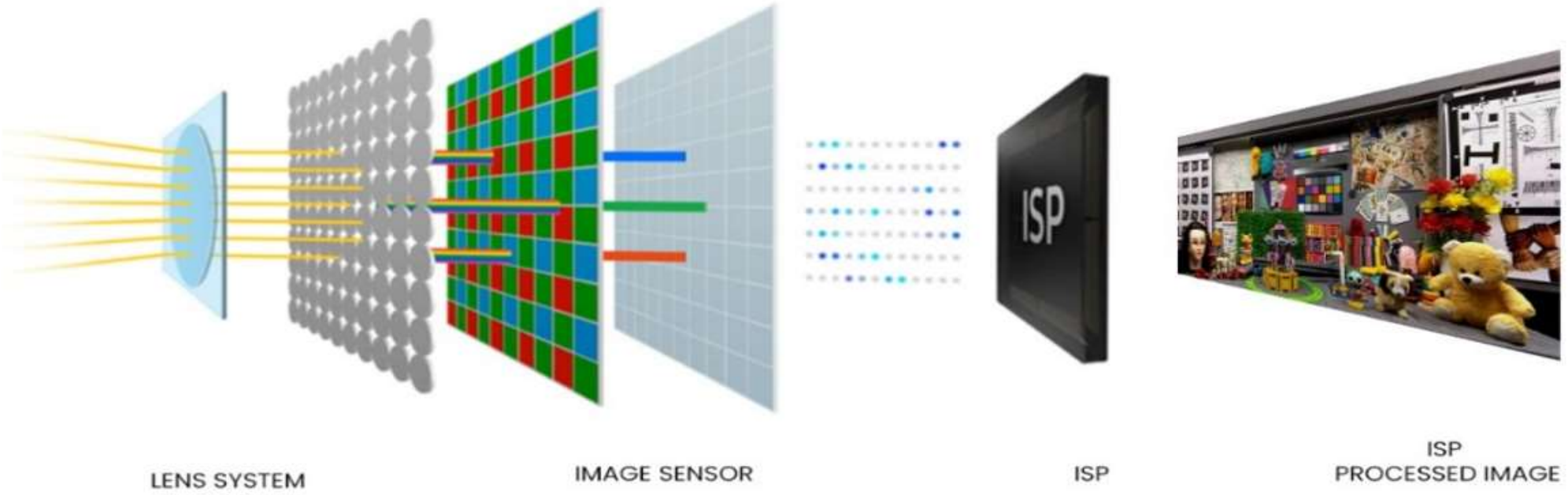
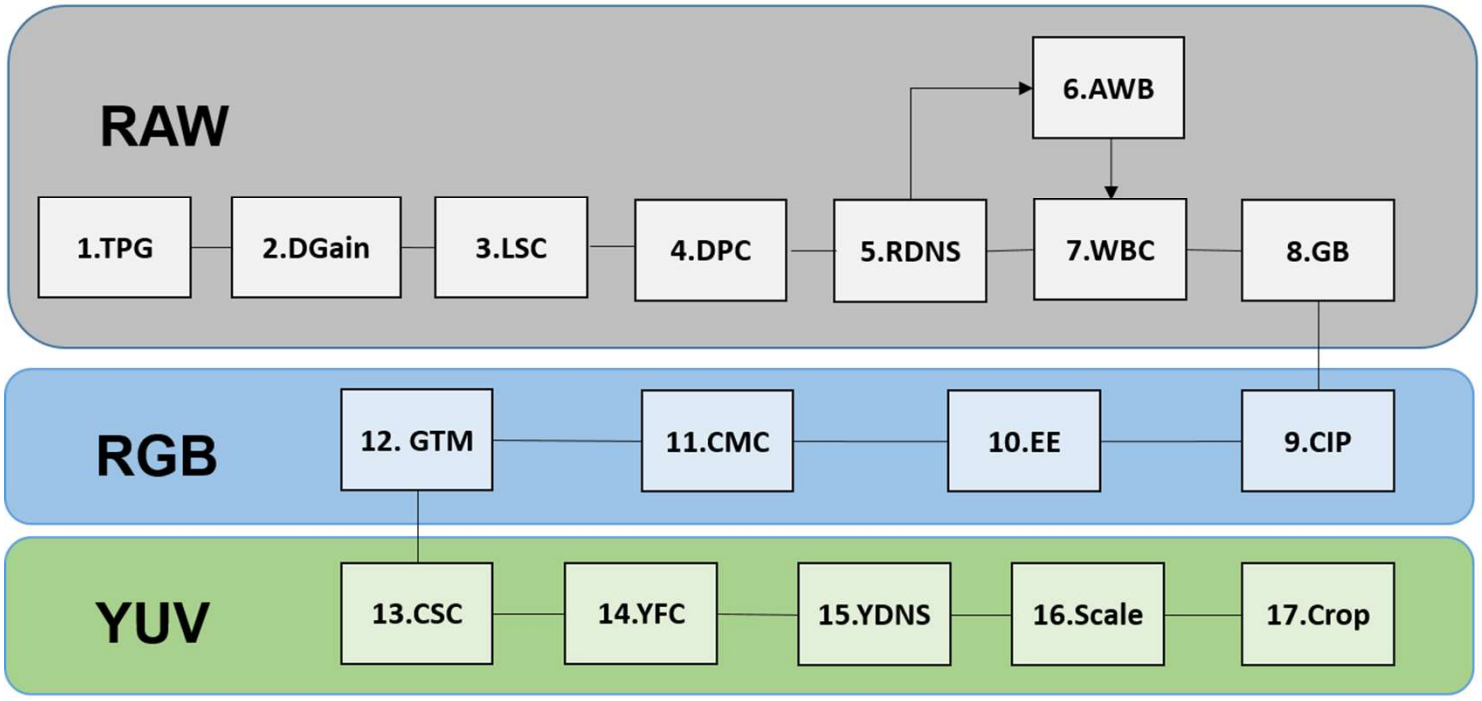


Image Signal Processor (ISP)



Source: PathPartner

Image Signal Processor (ISP)



xkISP: http://openasic.org

TPG	Test Pattern Generation
Dgain	Digital Gain
LSC	Lens Shading Correction
DPC	Defeated Pixel Correction
RAWDNS	RAW Denoise
AWB	Auto White Balance
WBC	White Balance Correctio
GB	Green Balance
DMC	Demosaic
EE	Edge Enhancement
CMC	Color Matrix Correction
GTM	Global Tone Mapping
CSC	Color Space Conversion
YFC	YUV Format Conversion
YUVDNS	YUV Denoise
Scale	Scaledown

Vision: new intelligence via deep learning