浙江大学计算机学院
数字媒体与网络技术

Digital Asset Management
数字媒体资源管理

2. Introduction to Digital Media Format

任课老师：张宏鑫

2008-09-05

Outline

- Image format and coding methods
- Audio format and coding methods
- Video format and coding methods
- Introduction to HTML and XML
- Graphics format and coding methods

Key points

- To grasp features of different types of digital media

- To understand principles of coding different types of digital media

Digital media data types

Table. File format used in Macromedia Director

File import					**File export**		**Native**
Image	Palette	Sound	Video	Animation	Image	Video	
BMP GIF, JPG, PICT, PNG, PNT, PSD, TGA, TIFF, WMF	PAL ACT	AIFF AU MP3 WAV	AVI MOV	DIR FLA FLC FLI GIF PPT	BMP	AVI MOV	DIR DXR EXE

2.1 Image format and coding methods

Common image formats

- GIF
- JPEG
- PNG
- TIFF
- TGA

- BMP (Win)
- PAINT&PICT (Mac)
- PPM (X-Win)
- WMF (Win)
- PS and PDF

- RAW
- DNG

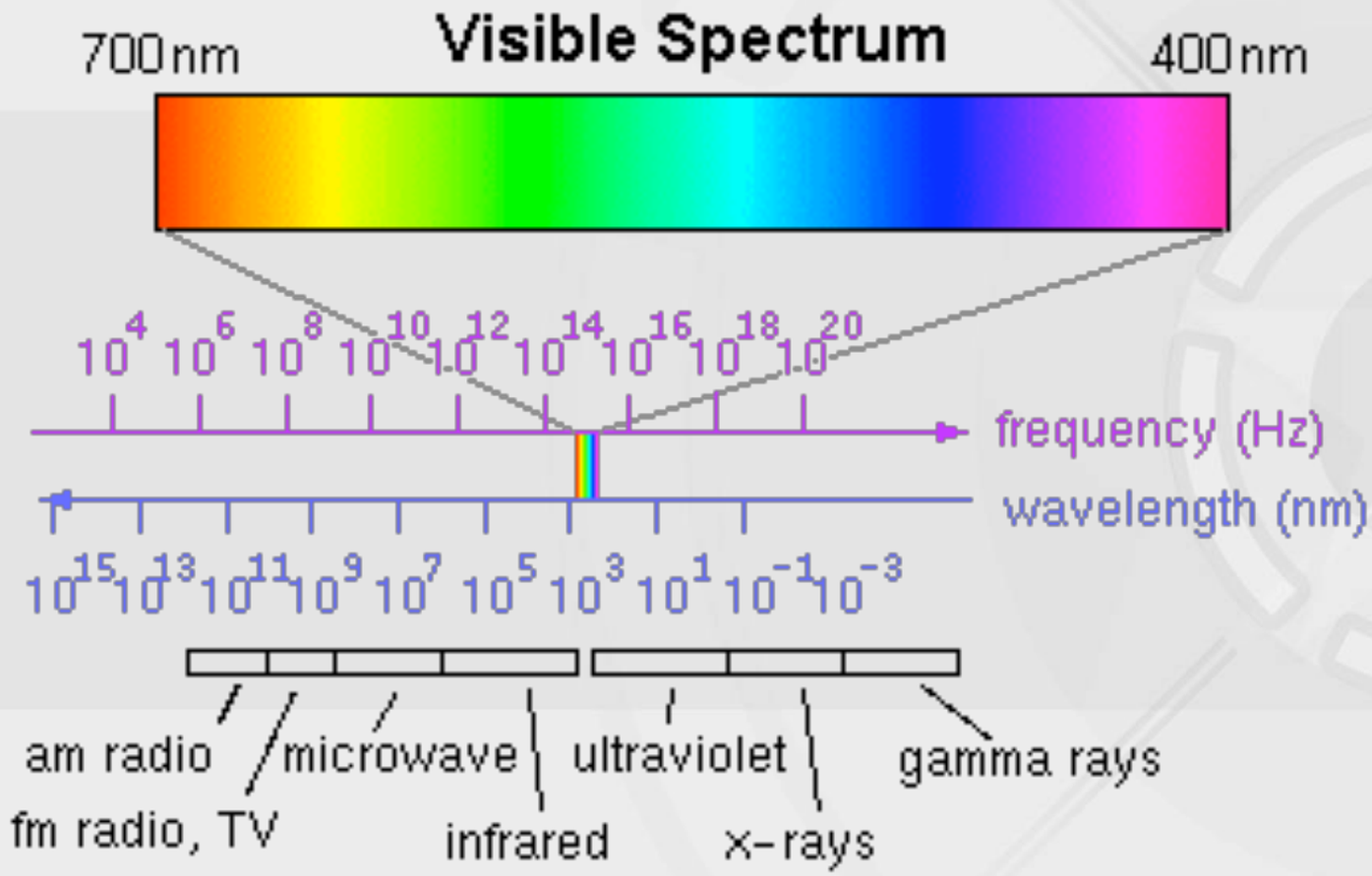
Common image formats

- Key points of storage
  - –Color space
  - –Coding methods
  - –Byte order： hardware dependent
    - MSB/LSB (most/least significant byte)

2.1.1 Color spaces

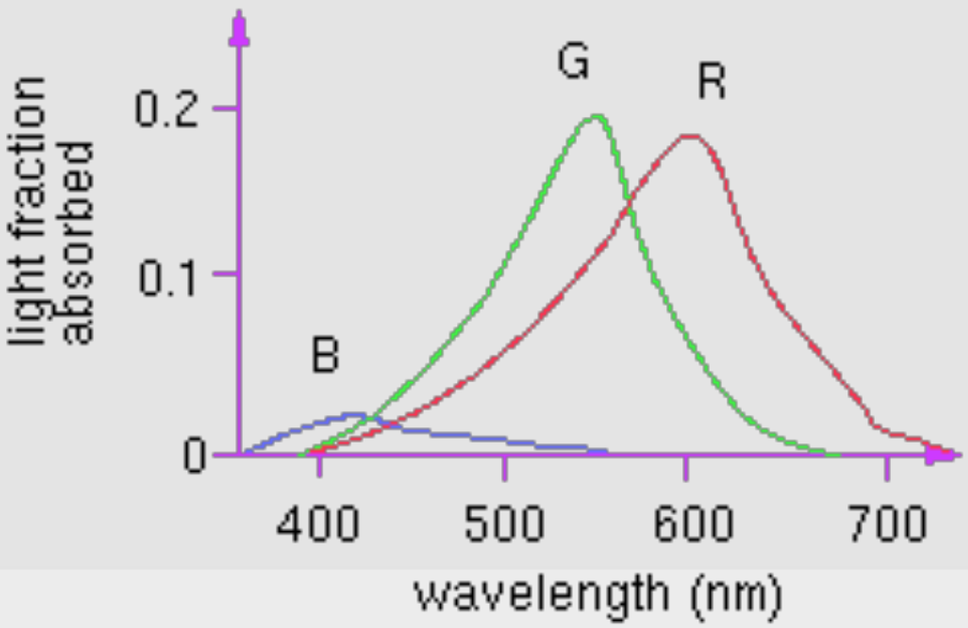
Color systems and color models

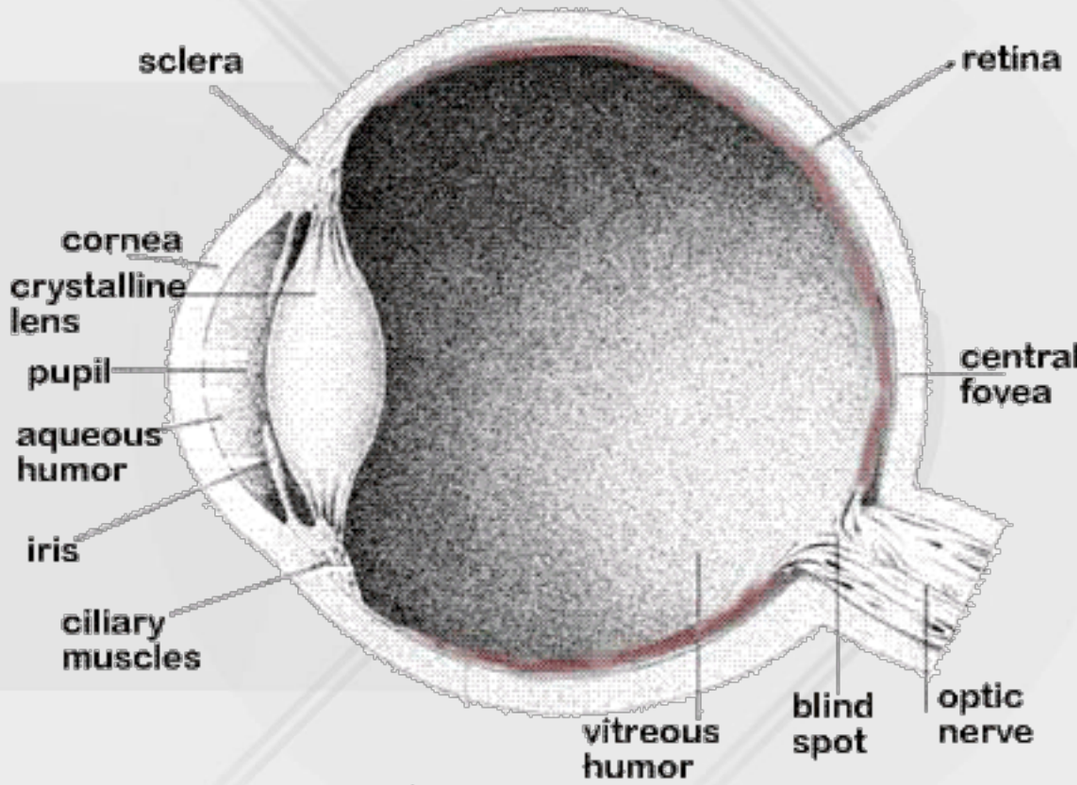
- RGB
- XYZ
- Lab
- YUV
- HSV

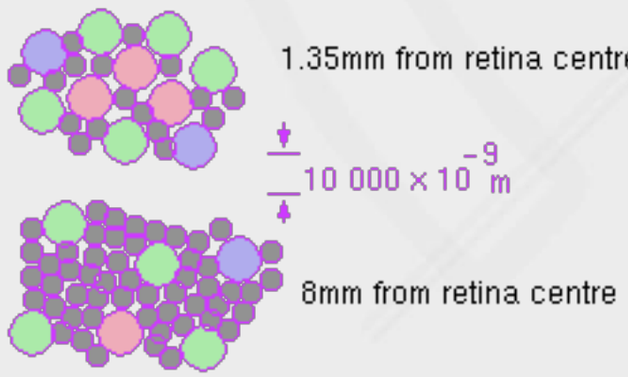


Color Model

The photosensitive part of the eye is called the *retina*. The retina is largely composed of two types of cells, called *rods* and *cones*. Only the cones are responsible for color perception. Cones are most densely packed within a region of the eye called the *fovea*.







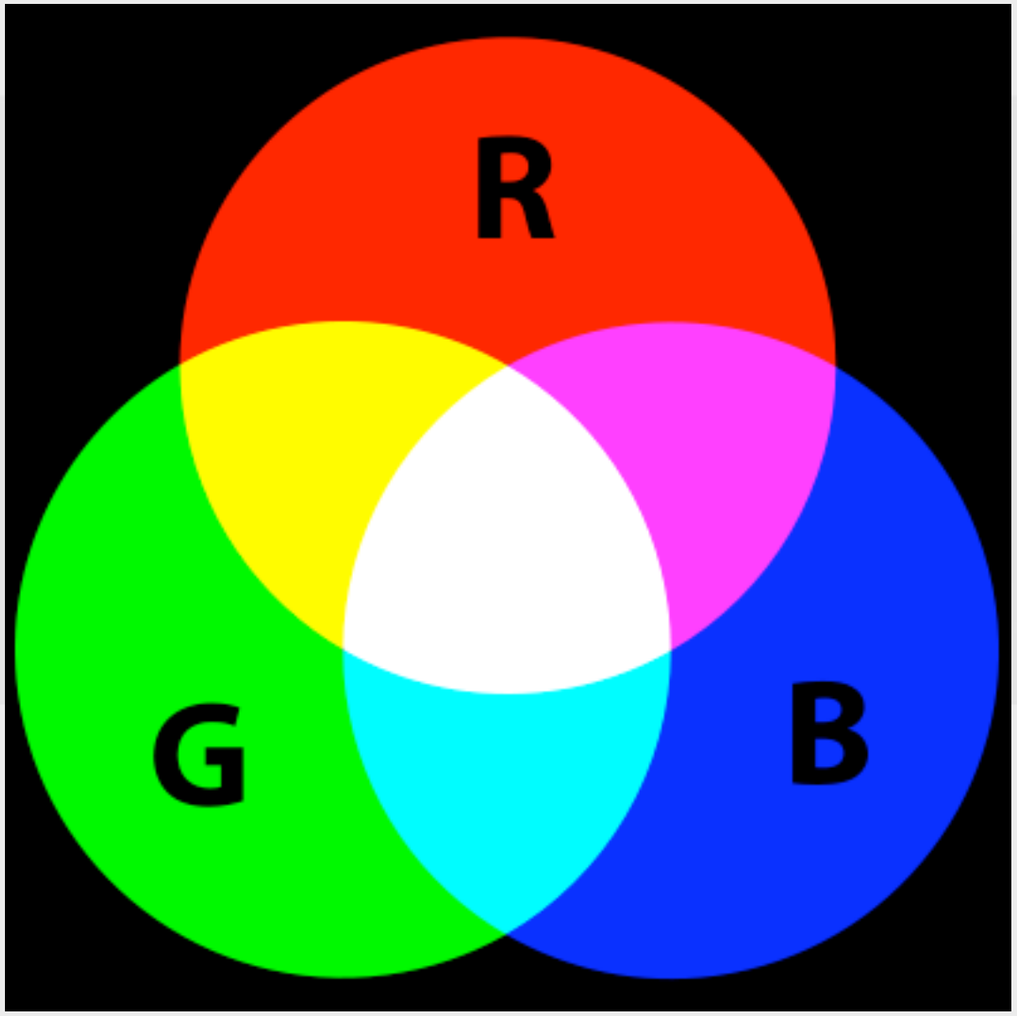
Color perception

- Three types of cones:

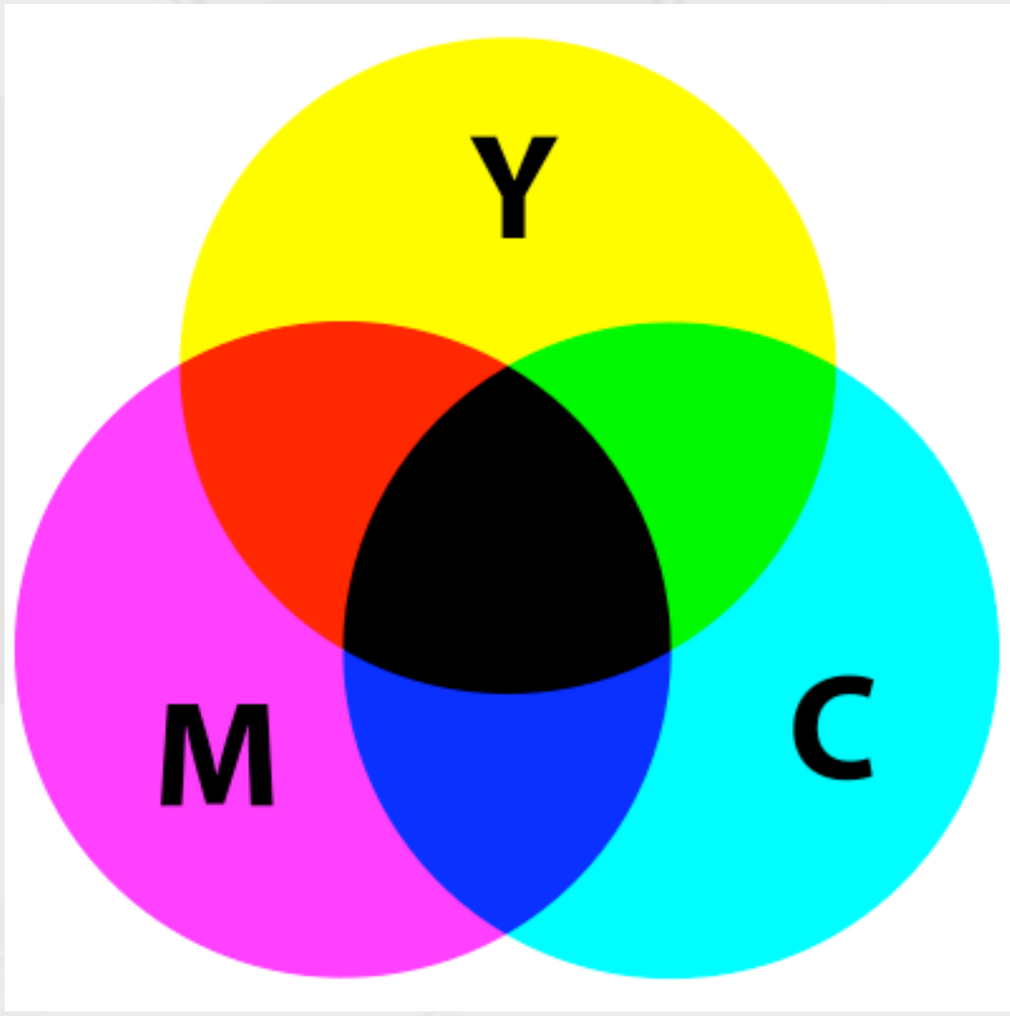
S	M	L	
Blue	Green	Red	roughly approximate
430nm	560nm	610nm	peak sensitivities

- Colorblindness results from a deficiency of one cone type.

RGB & CMYK



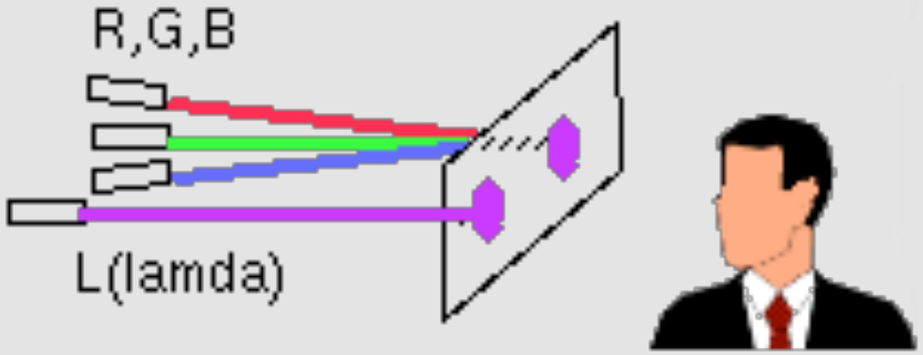
Additive color mixing

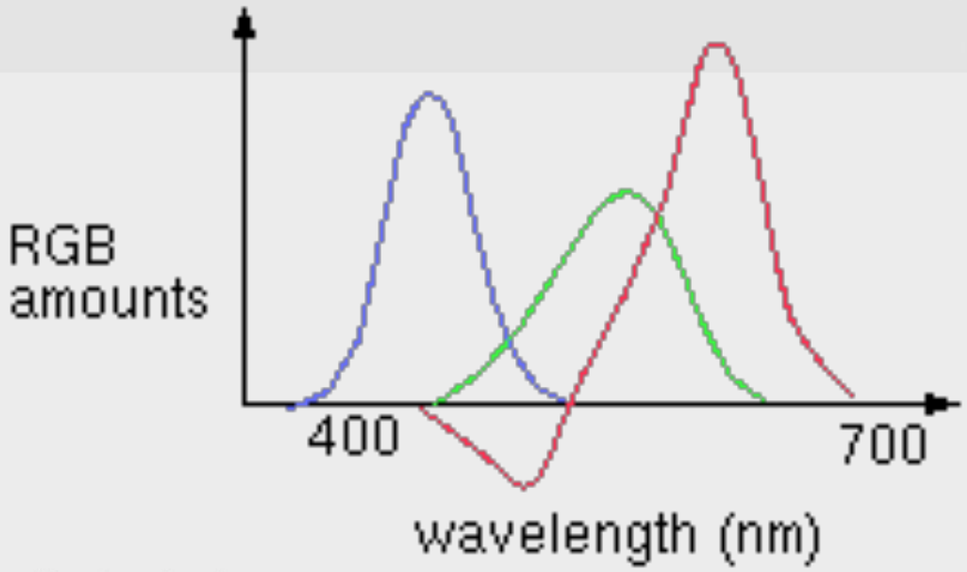


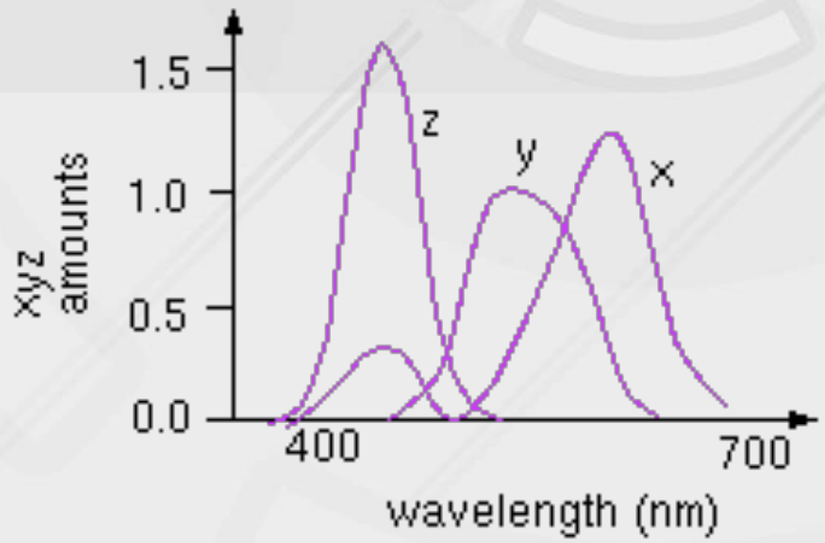
Subtractive color mixing

CIE XYZ space

- CIE: Commission Internationale d'Eclairage"







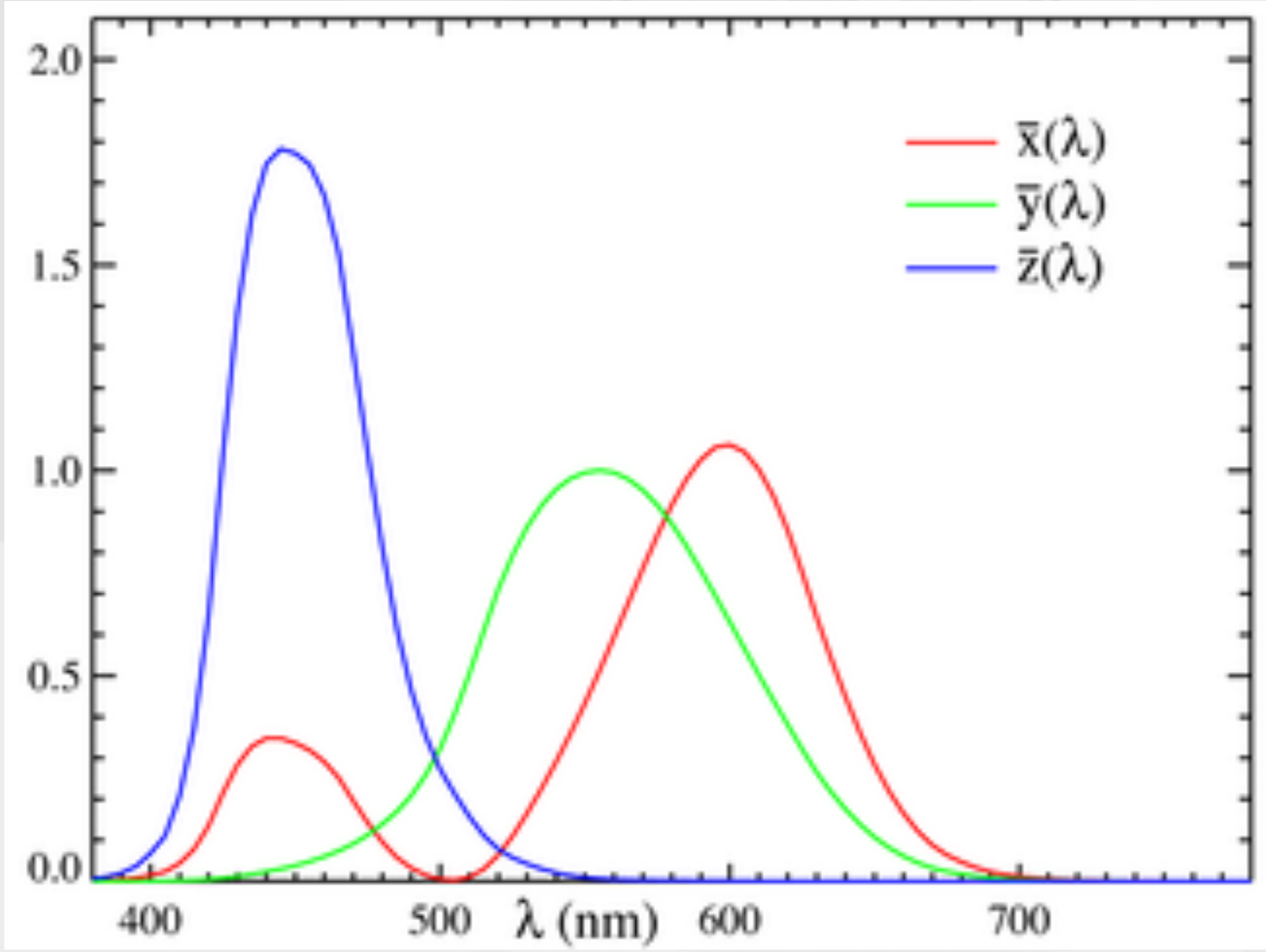
CIE XYZ space

- Color matching function

$$X = \int_0^{\infty} I(\lambda)\,\overline{x}(\lambda)\,d\lambda$$

$$Y = \int_0^{\infty} I(\lambda)\,\overline{y}(\lambda)\,d\lambda$$

$$Z = \int_0^{\infty} I(\lambda)\,\overline{z}(\lambda)\,d\lambda$$

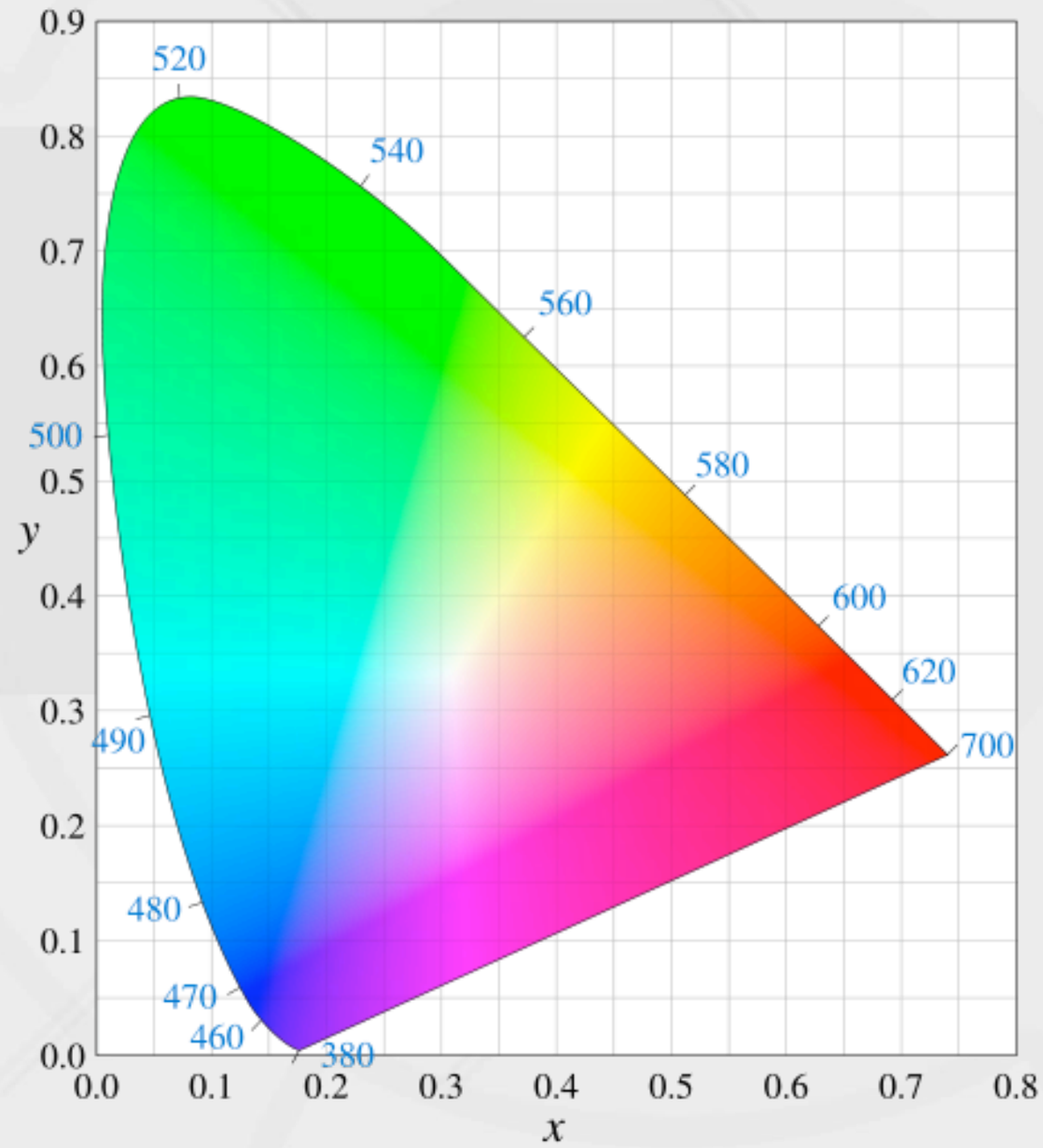


CIE XYZ space

$$x = \frac{X}{X+Y+Z}$$

$$y = \frac{Y}{X+Y+Z}$$

$$z = \frac{Z}{X+Y+Z} = 1 - x - y$$

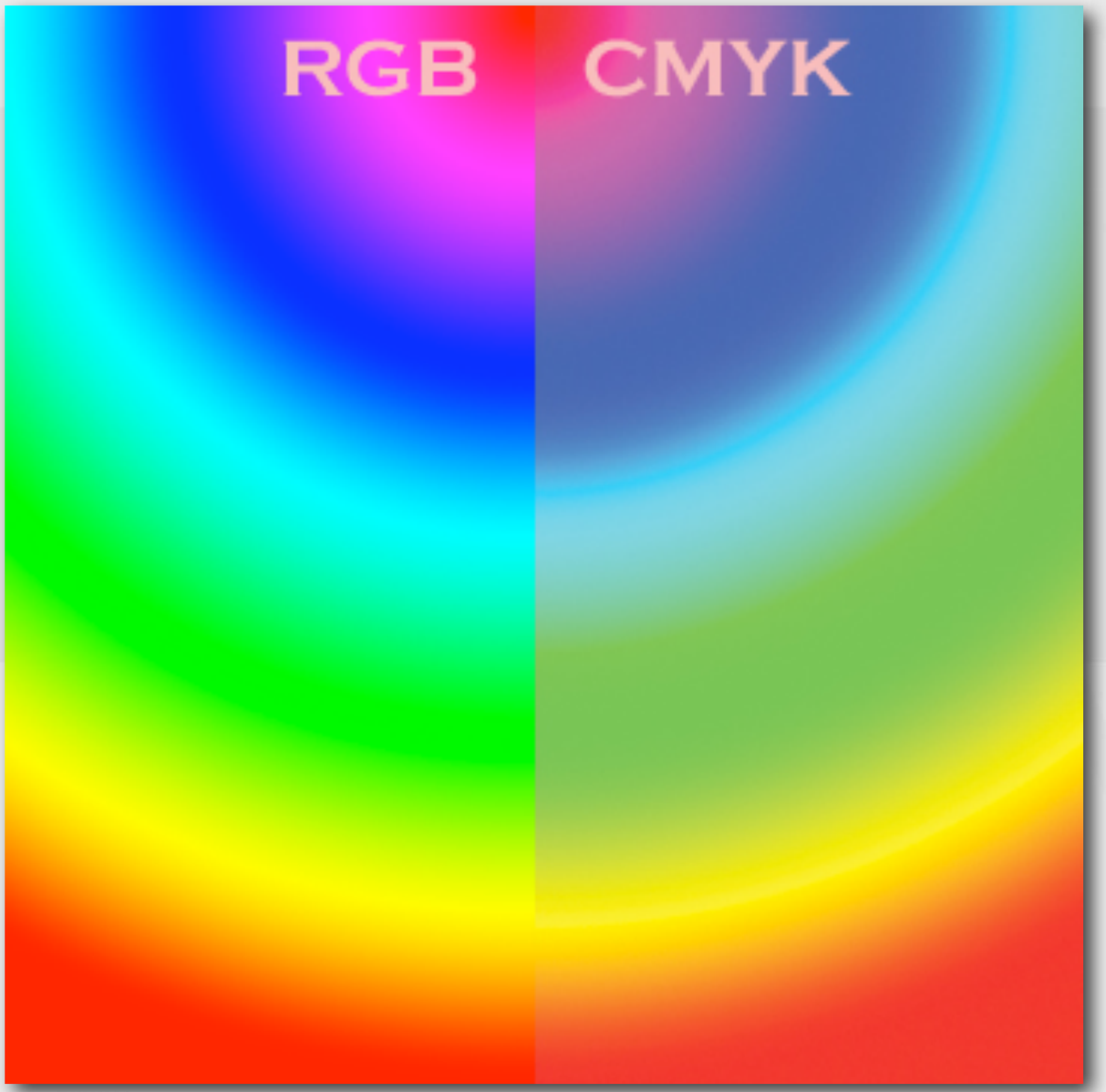


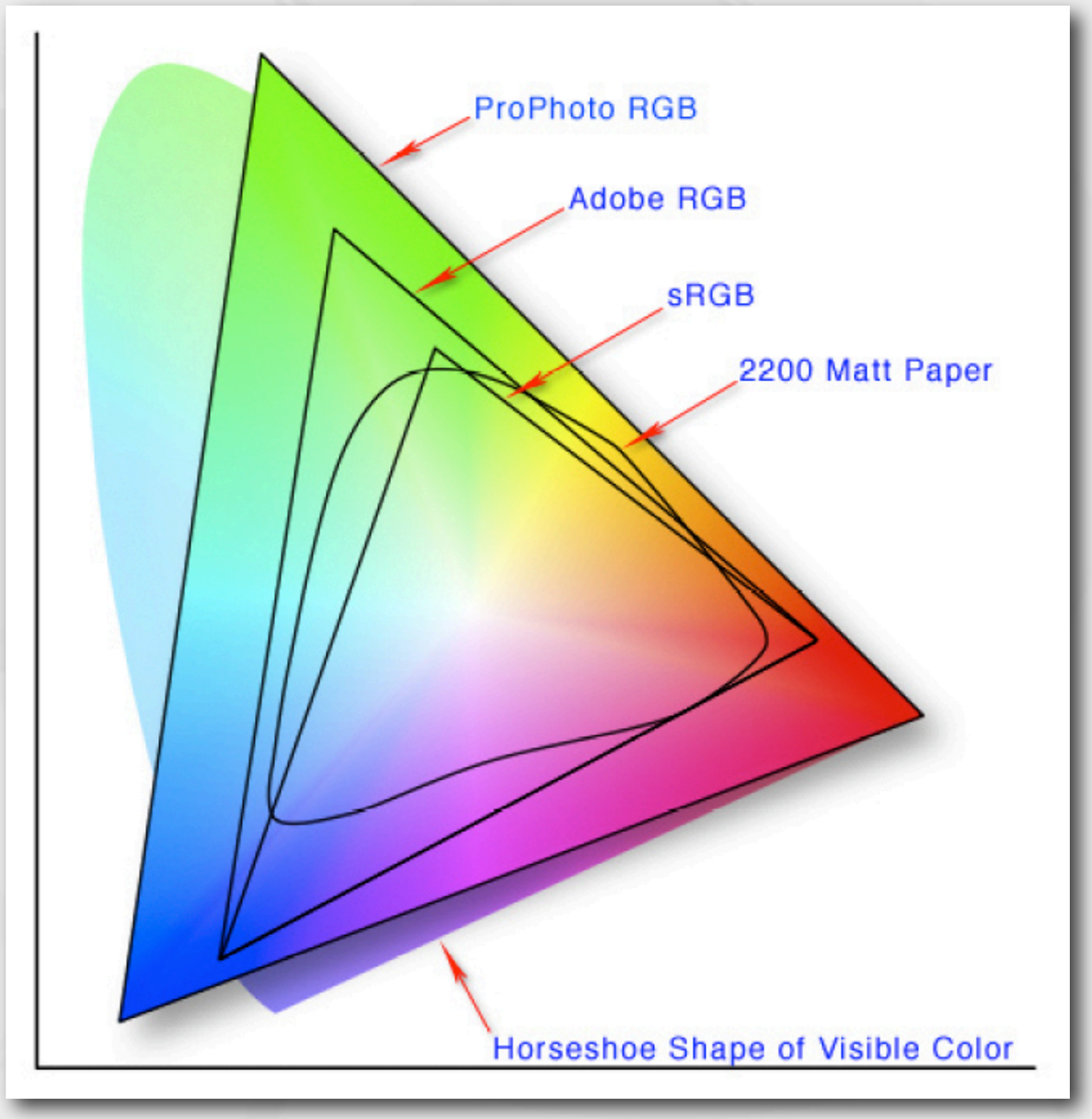
RGB vs. XYZ

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{b_{21}} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \frac{1}{0.17697} \begin{bmatrix} 0.49 & 0.31 & 0.20 \\ 0.17697 & 0.81240 & 0.01063 \\ 0.00 & 0.01 & 0.99 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Since 1931

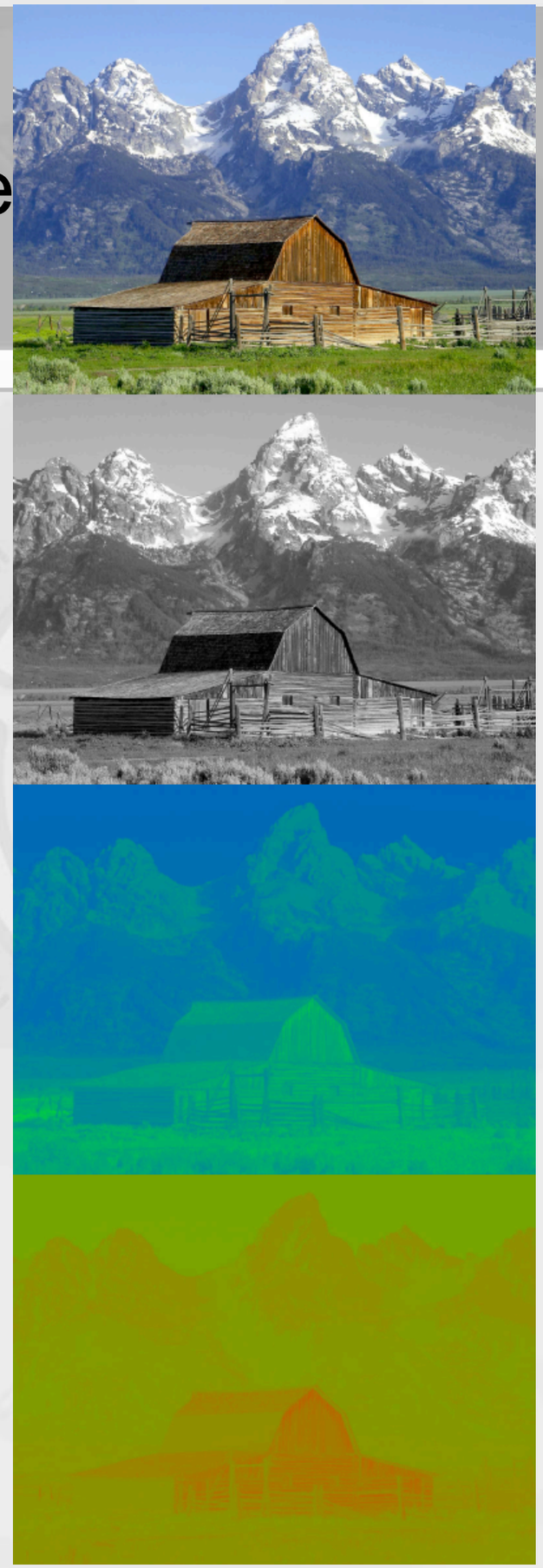
RGB :: CMYK :: XYZ color spaces





YUV color spaces

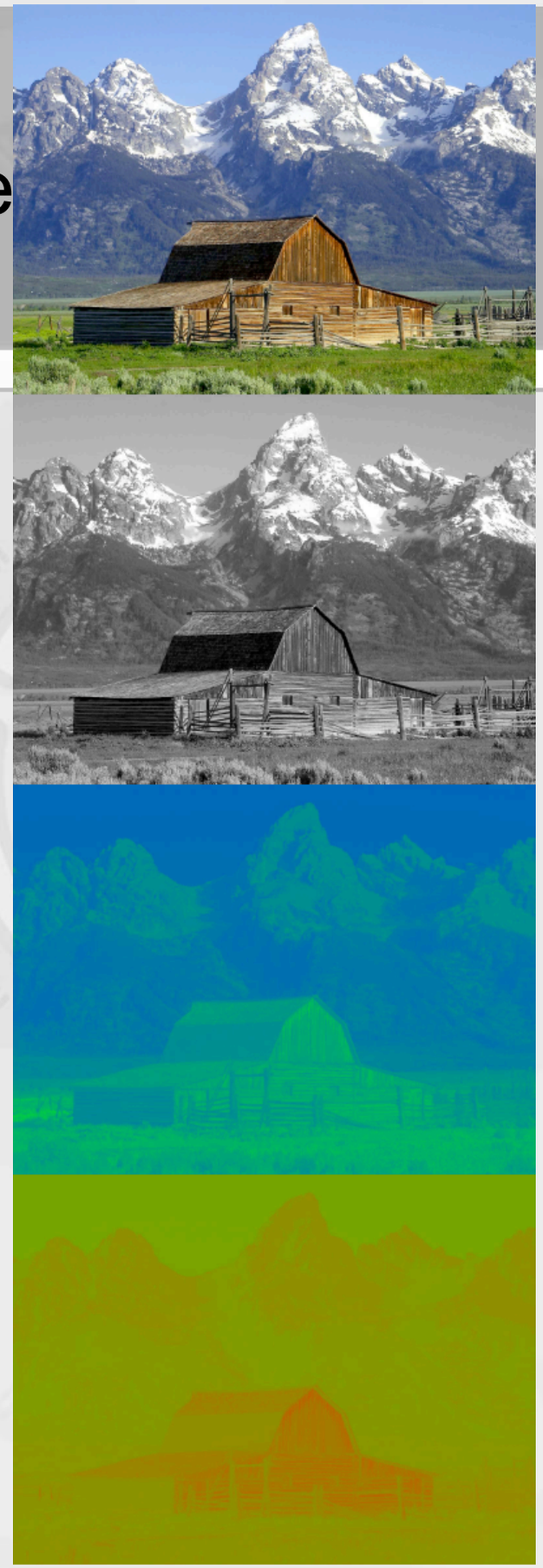
- used in most video capture system
- PAL television system



YUV color spaces

- used in most video capture system
- PAL television system

$$\begin{bmatrix} Y' \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.14713 & -0.28886 & 0.436 \\ 0.615 & -0.51499 & -0.10001 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$



Color spaces: reference

- http://en.wikipedia.org/wiki/Color_space

- http://www.cs.unc.edu/~mcmillan/comp136/Lecture4/Color.html

2.1.2 Image representations

Representation of digital images

- An image can be viewed as a ***N×M*** vector matrix
  - Grayscale image
  - Color image
  - Palette

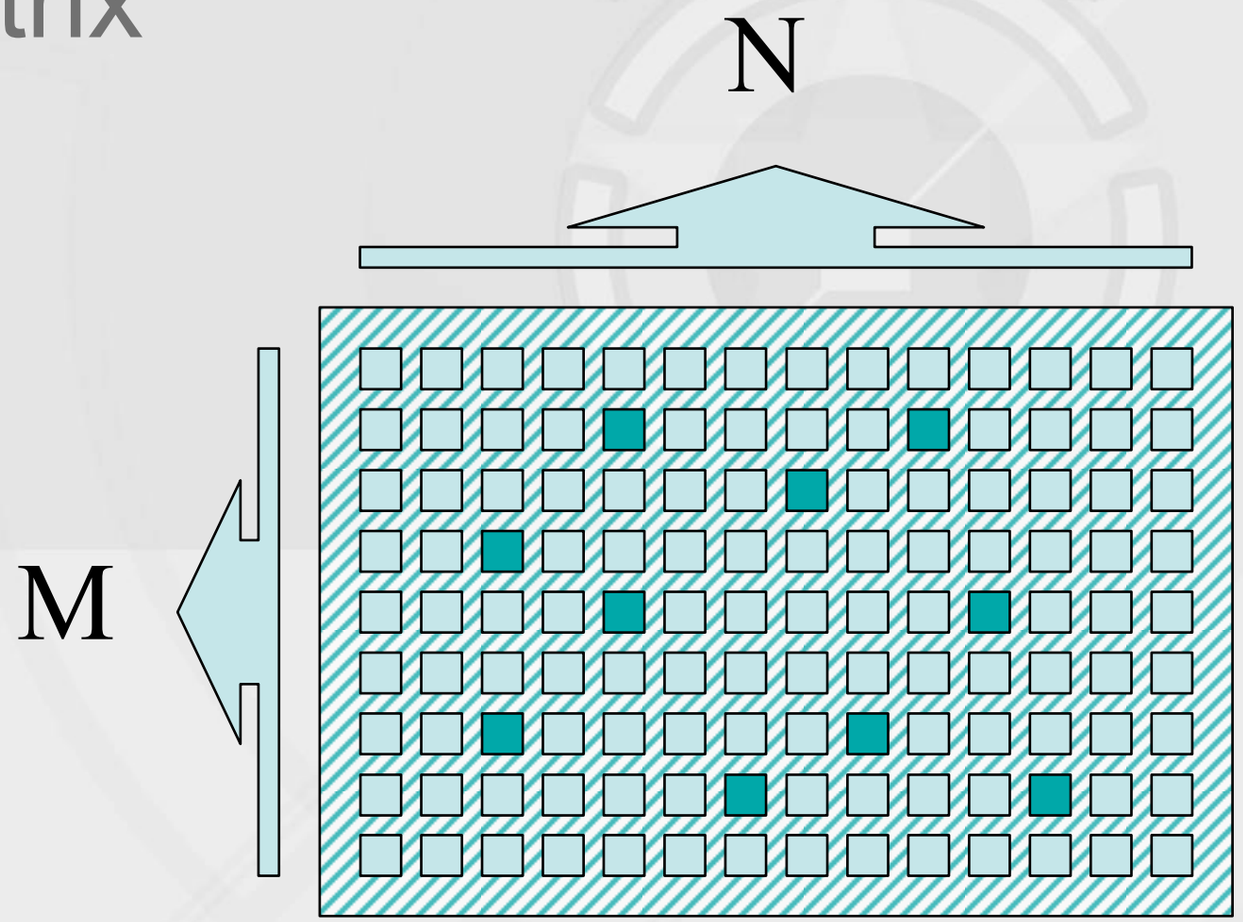
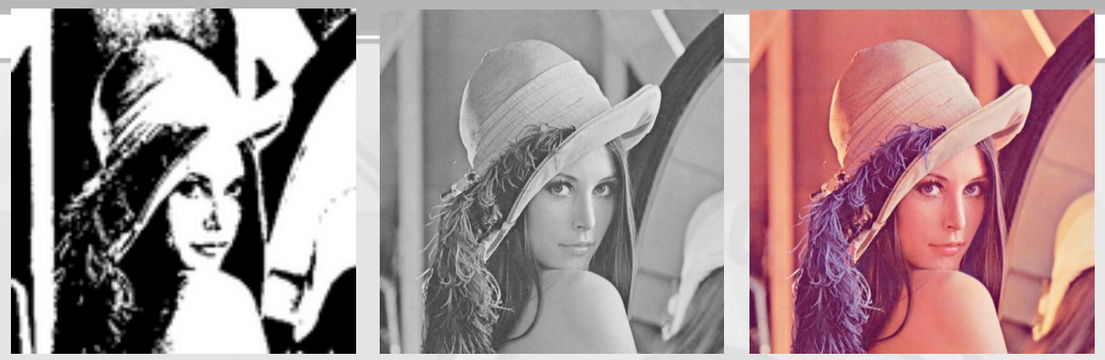


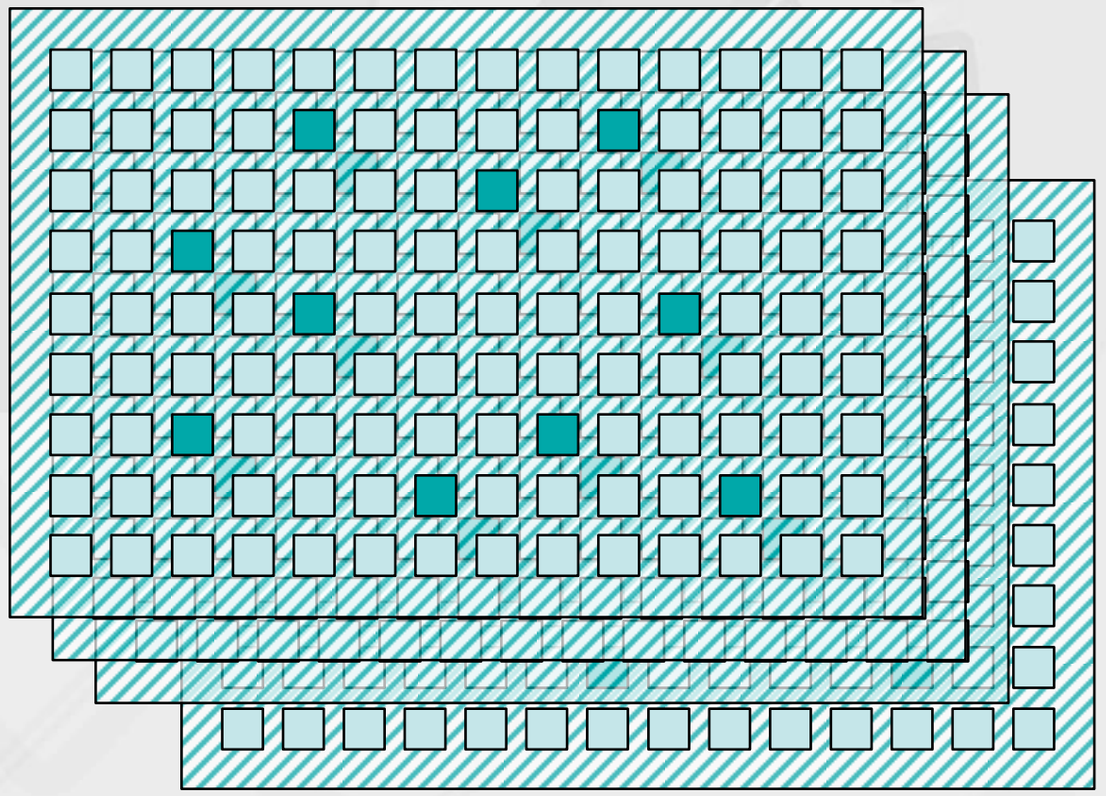
Image resolutions

Dimensions	MEGA pixels	Name	Comments
640x480	0.3	VGA	
720x576	0.4	CCIR 601 DV PAL	PAL DV, and PAL DVDs
768x576	0.4	CCIR 601 PAL full	PAL with square sampling grid ratio
800x600	0.4	SVGA	
1024x768	0.8	XGA	
1280x960	1.2		
1600x1200	2.1	UXGA	
1920x1080	2.1	1080 HDTV	high resolution digital TV format
2048x1536	3.1	2K	Typically used for digital effects in feature films.

Rep of Images

- Binary image
  - –1 bit = Boolean value
  - –One bit-planes



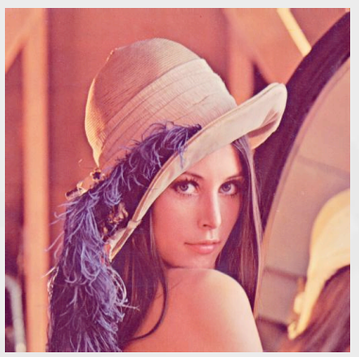


- Common Grayscale image:
  - –8 bits = 256 degrees of grayscale
  - –Eight bit-planes

Rep of Images

- Most used color images
  - –24bit RGB
  - –Red/Green/Blue each channel has 256 degrees of grayscale
  - –Can represent 2^24 = 16,777,216 types of color





Rep of image – Palette

- Some systems and applications can only use 8-bit color images
  - –Solution： Palette (Color look-up table)

High dynamic range image

HDRI example: a New York City nighttime cityscape

High dynamic range image

- HDR pixels:
  - 16-bit or 32-bit floating point numbers
  - 10–12 bits luminance, 8 bits chrominance
  - 10^−4 to 10^8: the range of visible luminance values

- CMOS image sensors: up to 110dB
- Tone mapping:
  - Typical computer monitors, prints, and other methods of displaying images only have a limited dynamic range

An example of a High Dynamic Range (HDR) photography, made of three different exposures

2.1.3 Image encoding

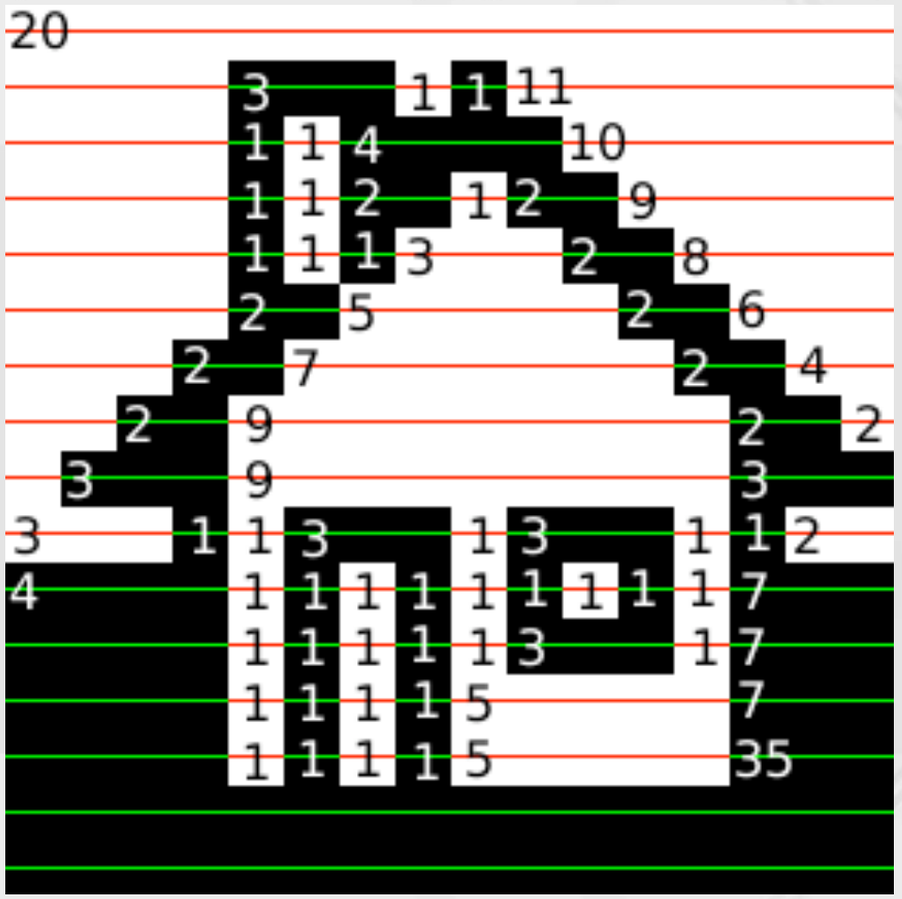
Image compression methods

- lossless compression
- lossy compression

Lossless image compression methods

- Based on information theory
- General encoding methods
  - –RLC (Run-Length Coding)
  - –VLC (Variable-Length Coding)
  - –Dictionary Coding
  - –Arithmetic Coding

Run-length Encoding



LZW : Lempel-Ziv-Welsh

- Universal lossless data compression algorithm
  - –by Abraham Lempel, Jacob Ziv, and Terry Welsh

- The compressor algorithm builds a string translation table from the text being compressed

LZW - Compressor

```
w = NIL;
add all possible charcodes to the dictionary
for (every character c in the uncompressed data) do
    if ((w + c) exists in the dictionary) then
        w = w + c;
    else
        add (w + c) to the dictionary;
        add the dictionary code for w to output;
        w = c;
    endif
done
add the dictionary code for w to output;
display output;
```

LZW - Decompressor

```
read a char k;
output k;
w = k;
while (read a char k) do
    if (index k exists in dictionary) then
        entry = dictionary entry for k;
    else if (k == currSizeDict)
        entry = w + w[0];
    else
        signal invalid code;
    endif
    output entry;
    add w+entry[0] to the dictionary;
    w = entry;
done
```

Deflate

- a lossless data compression algorithm:
  - –LZ77 algorithm + Huffman coding.
  - –originally defined by Phil Katz for version 2 of his PKZIP archiving tool,
  - –later specified in RFC 1951.
  - –used by gzip, modern versions of zip and as part of the compression process of PNG, PPP, HTTP, SSH

Lossless image compression methods (cont.)

- Other lossless image compression methods
  - –Image different encoding （差分）
  - –Lossless JPEG (JPEG 2000)
    - discrete wavelet transform

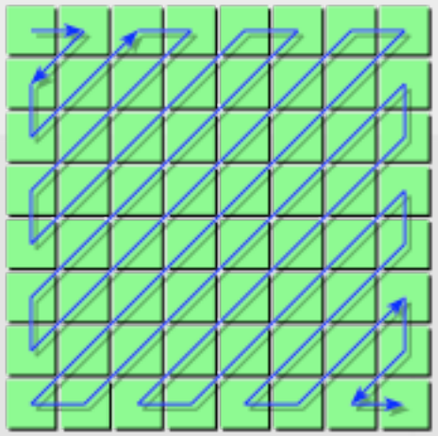
Lossy image compression methods

- Quantization
- Transform coding
  - –Discrete Cosine Transform
  - –Discrete Wavelet Transform
  - –Karhune-Loeve Transform (Principle component analysis)

Image compression standards

- JPEG
  - –Joint picture encoding group
  - –Discrete Cosine Transform
- JPEG 2000
  - –newer standard
  - –Discrete Wavelet Transform

JPEG compression: main idea



YCbCr Color Model

RGB Image

8x8 image blocks

DCT

Quantization

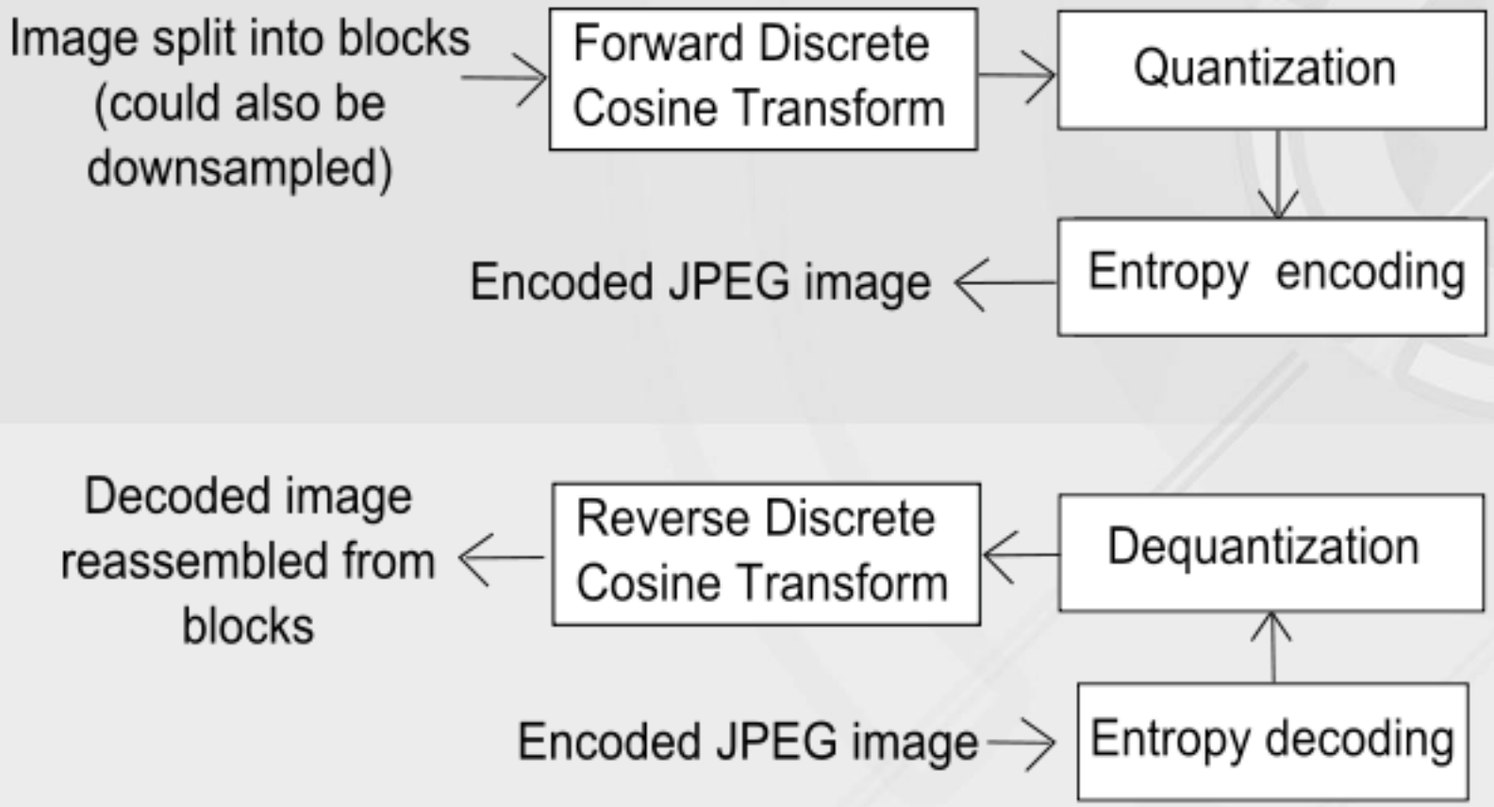
frequency-domain representation

$$\begin{bmatrix} 52 & 55 & 61 & 66 & 70 & 61 & 64 & 73 \\ 63 & 59 & 55 & 90 & 109 & 85 & 69 & 72 \\ 62 & 59 & 68 & 113 & 144 & 104 & 66 & 73 \\ 63 & 58 & 71 & 122 & 154 & 106 & 70 & 69 \\ 67 & 61 & 68 & 104 & 126 & 88 & 68 & 70 \\ 79 & 65 & 60 & 70 & 77 & 68 & 58 & 75 \\ 85 & 71 & 64 & 59 & 55 & 61 & 65 & 83 \\ 87 & 79 & 69 & 68 & 65 & 76 & 78 & 94 \end{bmatrix}$$

$$\begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

$$\begin{bmatrix} -26 & -3 & -6 & 2 & 2 & -1 & 0 & 0 \\ 0 & -2 & -4 & 1 & 1 & 0 & 0 & 0 \\ -3 & 1 & 5 & -1 & -1 & 0 & 0 & 0 \\ -4 & 1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

JPEG compression: implementation



Common image formats - GIF

- Graphics Interchange Format
  - UNISYS Corporation and Compuserve
  - Lempel-Ziv-Welch compression method
  - GIF87 / GIF89a

  - Features
    - Only support 8-bit (256) color image
    - Support several animation effects
    - Support interlaced image coding

Common image formats - PNG

- Portable Network Graphics
  - –motivation： Compuserv owns the LZW coding patent for GIF images
  - –open source
  - –Transparent
  - –PNG64

Common image formats - JPEG

- Lossy to lossless editing

DNG: Digital Negative

- a royalty free RAW image format
- design by Adobe
- based on TIFF/EP
- mandates use of metadata

Common image formats - TIFF (6.0)

- Tagged Image File Format
  - –flexible and adaptable
  - –handling images and data within a single file
  - –header tags: size, definition, image-data arrangement, applied image compression
  - –defining the image's geometry.

Common image formats - TIFF (6.0)

- a TIFF can be a container file
  - –compressed JPEG and RLE
  - –lossless compression

- include a vector-based Clipping path (outlines, cropping, image frames)

Summary – Essential factors of image storage

- Resolution
- Compression rate
  - –1bpp,2bpp, …
  - –Compression methods
- Color representation
  - –RGB, YUV, Lab …

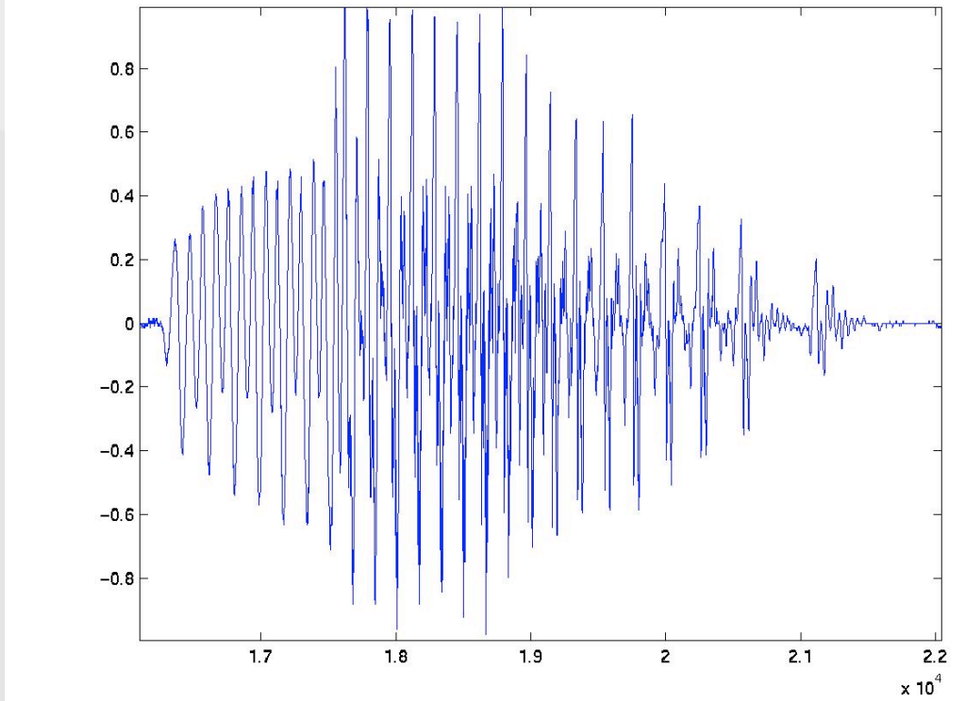
Image converting tools

- ACDSEE
- XnView
  - http://perso.orange.fr/pierre.g/

2.2. Audio formats and compression methods

Digitalized audio / sound

- What is sound?
  - –Knowing from ear?!?
  - –Sound wave ?!?



- Digitalization
  - –Analog signal → digital signal
  - –Quantization

Bit rate and bit

- a kind of energy wave.
- a continuous function of wave amplitude
  - Sequence is related to the X axis (the time line).
  - Amplitude is related to the Y axis.

Higher coding rate and quantization rate, better sampling quality

- discretely sampled during the digital coding period
  - Bit rate: number of samples obtained in one second
    - The highest frequency ~ 20kHz.
    - 40k samples per second (Nyquest theorem)
    - The bit rate of CD is 44.1kHz
  - Quantization rate: must be the power of 2.
    - The quantization rate of audio CD is normally 16bit.

Audio compression: lossless vs. lossy

- There is no absolute looseness coding schemes!
  - According to the definitions of bit rate and quantization rate, audio coding can only approximate to the natural sound signal as much as possible.
  - Comparing with natural signal, all coding schemes are lossy.

- Related looseness scheme: PCM
  - PCM can reach the highest preserving level.
  - widely applied in raw data saving and music data, e.g. CD、DVD and WAV files.
  - PCM is viewed as a looseness coding scheme. How, PCM only approximate to the raw data.
  - Comparing with the PCM coding method，we usually put MP3 coding methods into the lossy audio encoding methods.

PCM coding

- PCM - **P**ulse **C**ode **M**odulation

- PCM coding
  - Advantage：good play back quality.
  - Shortage: large storage space.

- Audio CD mainly leverage the PCM coding scheme. One piece of CD can store 72 minutes music.

PCM audio stream bit-rate

- Formula
  - Bit rate × Quantization rate × number of sound channels (bps)。

- EXAMPLE：
- WAV file: bit rate 44.1KHz，quantization rate 16bit，stereo sound.
  - Coding rate: 44.1K×16×2 =1411.2 Kbps.
  - 128K MP3 ~ 1411.2 K bits per second
  - also called data width, similar to the concept of band width used in network transfer.

  - Data speed: transferred bytes per second, = Bit rate / 8. In this example, the speed is 176.4KB/s.
  - It takes space of 176.4KB per second. Recording 1 minute music requires 10.34M.

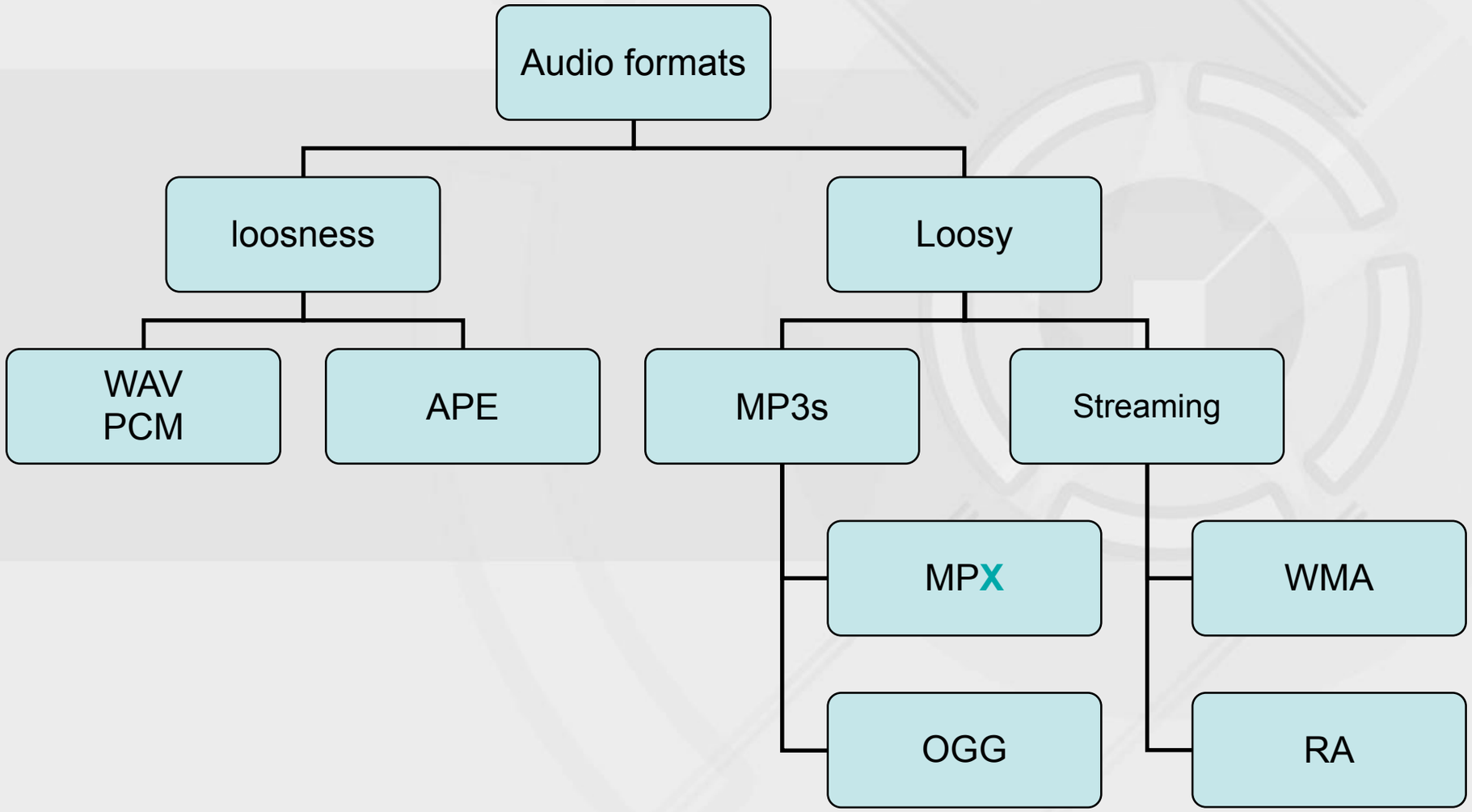
The streaming feature of audio

- The blooming of network => play on-line music.
  - –play the music meanwhile downloading.
    - Recent techniques are easy to archive this goal.

- Based on this feature, it is easy to implement:
  - –on-line direct-show
  - –DIY digital broad casting.

Common audio formats

- WAV
- MP3
- WMA
- RA
- OGG
- APE



WAV

- Developed by Microsoft

- WAV format is based on PIFF Resource Interchange File Format standard.
  - All WAV files have a file head which is used to record coding parameters of audio stream. WAV file have no specific constraints on coding audio stream. Besides PCM, WAV can use any types of coding schemes defined by ACM.

- In Windows，PCM based WAV format is recognized as a most useful audio format. WAV is good for music creating and editing, and for saving raw music data.
  - PCM based WAV file is now employed as an intermediate format for convert over different type of audio data, e.g., MP3 to WMA.

WMA

- WMA is created the Windows Media Audio coding framework, developed by Microsoft.
- WMA is designed to used for network transfer. Its main competitors are products from Real Networks.
  - Microsoft claimed that WMA can reach the sound quality of CD in 64kbps bit rate.
  - Provides Windows Media Rights Manager to prevent illegal copies and to count play times.
  - Supports stream techniques and online broadcasting.

RA

- RA (RealAudio) is proposed by RealNetworks Inc.
- In network application, many music site use RealAudio for online playing.

- RA mainly focus on network media market
  - Highlight: RA can alter its own coding bit rate due to the network width but keep the sound quality as much as possible.
  - RA can support many types of audio coding schemes, e.g., ATRAC3.
  - Beside the function of download-while-play, RA can also hide true internet address of sound file. It is quite useful for Music company

APE

- APE is a looseness compression format proposed by Monkey‘s Audio.
- They mainly used LZW as the compression kernel.
- High compression ratio but fast compression speed.
  - Used by many music fans to record CD and share music resources.
- Monkey‘s Audio provides a set plug-ins for different types of media players.

OGG

- OGG is a huge project plan of multimedia R&D and is mainly focus on video/audio coding.
  - The total OGG project is open source and free

- Ogg Vorbis audio coding
  - Comparing with MP3, it provides lower bit rate but better play back quality.
  - Support more channels than MP3. It is suitable for recoding classical music.
  - Flexible audio coding framework

MP3

- From the MPEG-3 standard

什么是**MIDI**

- MIDI (Musical Instrument Digital Interface即乐器数字化接口) is an international standard for general interface.
  - It provides a set of standard interface for transferring data among different types of devices. MIDI devices shall precisely send MIDI messages.

- Wildly use in music creation, game background music and ring tone of mobile phones.

MIDI概况

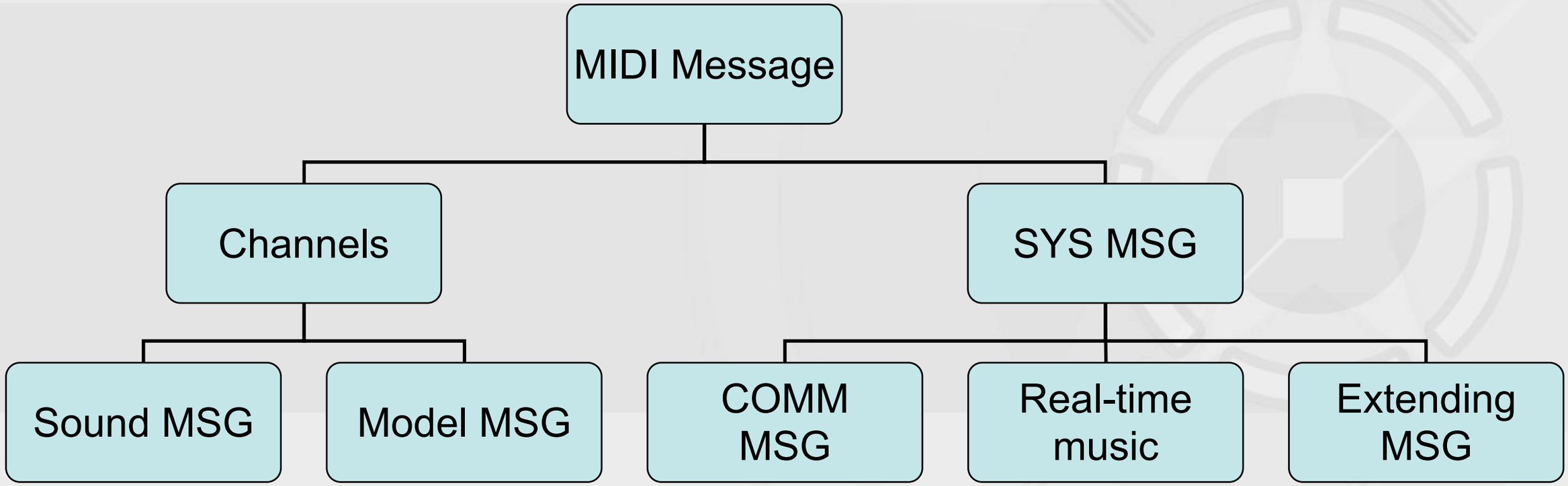
- MIDI is type of description language.
  - Different directly record digitalized sound signal
  - Only record ‘events’ that how instruments make sound.
  - Small storage size.

- Three elements of MIDI
  - Synthesizer
    - Generate sound and can control the length, height, strength and other features of sound.
  - Sequencer
    - Devices or software that store and modify MIDI information.
  - MIDI device
    - Do not generate any sound but a sequence of MIDI commands.
    - E.g. MIDI keyboard, MIDI harp, MIDI guitar, and MIDI violin, etc.

Basic concepts of MIDI

- [Track]
  - Music is composed with several music channels.
- [Channel]
  - Each MIDI device corresponds to a channels. Each channel owns its own message sequence. Up to 16 channels
- [Voice]
  - Each channel allows multiple voice, e.g., chords when playing piano. (*Timbre* means the sum of sound in one channels)
- [Polyphony]
  - The sum of sound can be generated by Synthesizer in one moment.
- [Patch]
  - Sound feature setting up to simulate specific instrument.

Message structure of MIDI



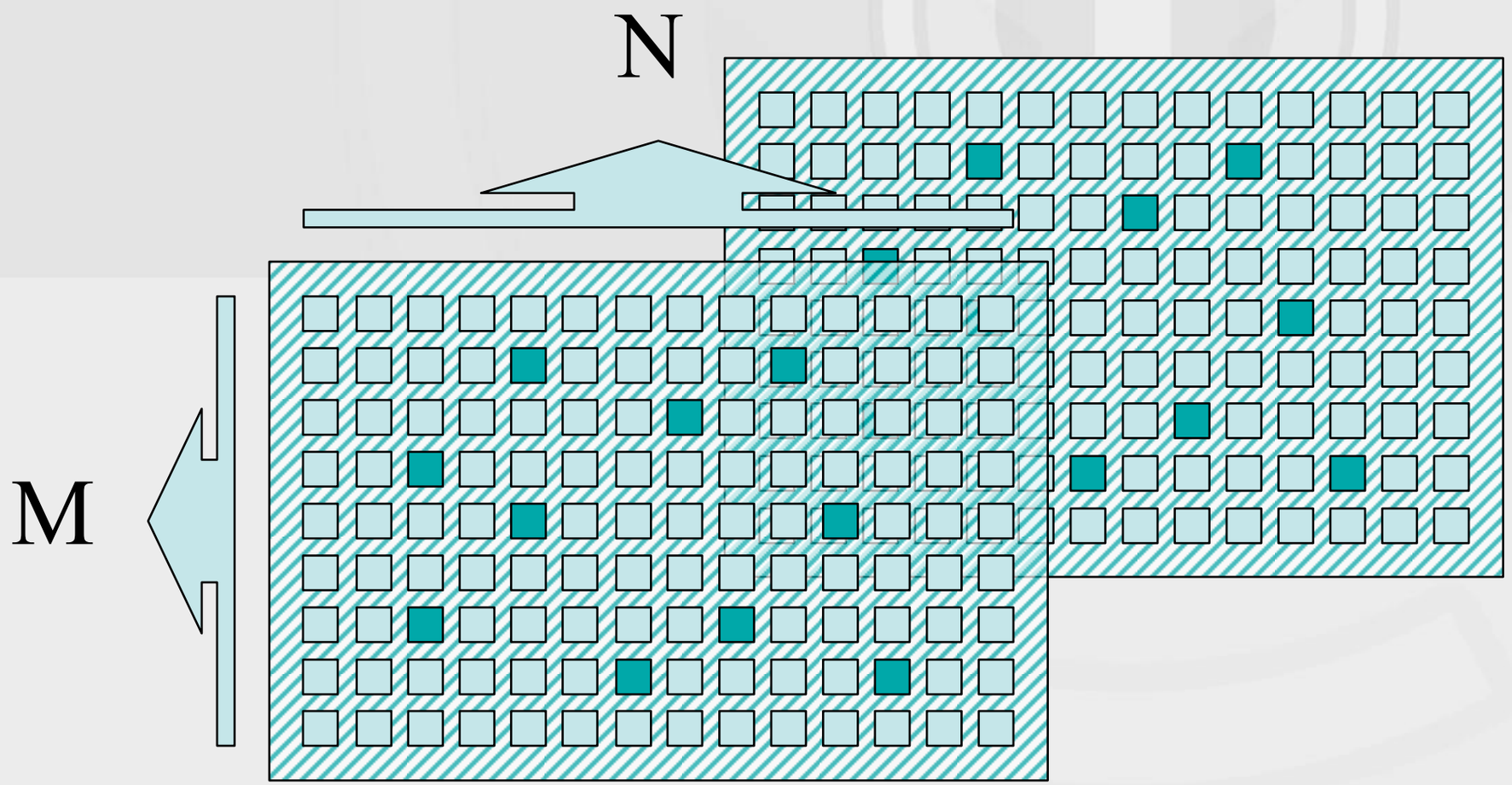
Common MIDI file format

- MID
  - –General MIDI
- SMF
  - –Standard MIDI File

2.3. Video formats and coding methods

Representations of video

- Sequence of images ？！？
  - –Can be viewed as a 3-dimentional matrix
  - –But it is only 50% correct



Common video formats

- AVI (Microsoft, Divx, …)
  - –avi, wmv, asf
- RM (Realplayer)
  - –rm, rmvb
- MOV (Quicktime)
  - –mov
- MPEG
  - –MPEG-1, MPEG-2, MPEG-4 …

http://www.bigbuckbunny.org/index.php/download/

Common video formats - AVI

- AVI = Audio Video Interleaved (By Microsoft)
  - A digital audio/video format according to the RIFF file format standard.
  - multimedia CDROM, store video information, movie and TV program,
  - Internet applications, download and online viewing

- Allows storing audio and video information interlaced
- But play back simultaneously

Common video formats - AVI

- AVI only defines the standard on control interface.
  - –No limitation of compression approach in AVI file format
  - –Supports 256 colors and RLE compression
  - –AVI with specific encoding methods must be played back by matched decoding methods.
  - –Many companies provide their own codecs
    - e.g., SONY

Common video formats - RM

- RM (RealVideo file): a new file format for streaming video by RealNetworks Inc.

- RealVideo techniques is used to broadcast important events over Internet.

- RealMedia: A audio/video compression standard of RealNetworks
  - Mainly used in wide range network to transform real-time video sequence in low bit rate.
  - It can alter different bit rate depends on network data transformation rate

- RealVideo can be used with RealServer. Different from most other video formats, RM can be played back while the data is downloading.

Common video formats - MOV

- A video/audio format developed by Apple Inc.

- QuickTime™ player
  - –Apple Mac OS、Microsoft Windows System
- The original format supports
  - –256 color, RLE, and JPEG compression techniques.

Common video formats - MOV

- Advanced function features
  - > 150 kinds of Video effects
  - > 200 kinds of MIDI devices sounds.

- Internet-oriented features
  - –digitalized information stream,
  - –workflow, and
  - –play-back functions through internet.

Common video formats - MOV

- QuickTime VR (QTVR):
  - –a set of Virtual Reality (虚拟现实) techniques used in QuickTime.
  - –use mouse or keyboard
    - investigate 360 degree of scene
    - browse an object from a specific spatial angle interactively.

Video compression standards

- MPEG standards
  - Audio/Video compression, storage and play back standards
  - MPEG-1: VCD
  - MPEG-2: broadcast TV, e.g., DVD、HDTV etc.
  - MPEG-3: replaced by MPEG-2
  - MPEG-4: network video transfer，stream media
  - MPEG-7:
  - MPEG-21:

- ITU-T H.26x series

Video compression standards

- ITU-T H.26x series
  - Mainly used in video communication applications
  - Now it has H.261，H.262，H.263，H.264

  - ISDN network based H.320 standards
    - the video compression part: H.261, H.262 and H.263

  - LAN network based H.323
  - PSTN network based H.324
    - the video compression part: H.261 and H.263

MPEG概况

- MPEG = **M**otion **P**icture **E**xpert **G**roup

- **ISO/IEC JTC1/SC29**
  - WG11:Motion Picture Experts Group (MPEG)
  - WG10:Joint Photographic Experts Group (JPEG)
  - WG7: Computer Graphics Experts Group (CGEG)
  - WG9: Joint Bi-level Image coding experts Group (JBIG)
  - WG12: Multimedia and Hypermedia information coding Experts Group (MHEG)

MPEG概况

- **MPEG-1,2 standards were started at 1988**
  - 需求 **[Requirement]**
  - 系统 **[System]**
  - 视频 **[Video]**
  - 音频 **[Audio]**
  - 实现 **[Implementation]**
  - 测试 **[Testing]**

- **Newest MPEG standards: MPEG-4, MPEG-7, MPEG-21**

MPEG-1 Standard ISO/IEC 11172-2 (1991)

"Coding of moving pictures and associated audio for digital storage media"

- **Video**
  - optimized for bit rates around 1.5 Mbit/s
  - originally optimized for SIF picture format,
  - but not limited to it:
    - [**NTSC based**] : 352x240 pixels at 30 frames/sec
    - [**PAL based**] : 352x288 pixels at 25 frames/sec
  - progressive frames only
    - no direct provision for interlaced video applications, such as broadcast television

MPEG-1 Standard ISO/IEC 11172-2 (1991)

- Audio
  - joint stereo audio coding at 192 kbit/s (layer 2)

- System
  - mainly designed for error-free digital storage media
  - multiplexing of audio, video and data

- Applications
  - CD-I, digital multimedia, and
  - video database (e.g. video-on-demand)

MPEG-2 Standard ISO/IEC 13818-2 (1994)

- Video
  - 2-15 or 16-80 Mbit/s bit rate (target bit rate: 4…9 Mbit/sec)
  - TV and HDTV picture formats
  - Supports interlaced material

  - MPEG-2 consists of *profiles* and *levels*
    - Main Profile, Main Level (MP@ML)
      - 720x480 resolution video at 30 frames/sec
      - < 15 Mbit/sec (typical ~4 Mbit/sec)
      - for NTSC video

    - Main Profile, High Level (MP@HL)
      - 1920x1152 resolution video at 30 frames/sec
      - < 80 Mbit/sec (typical ~15 Mbit/sec)
      - HDTV

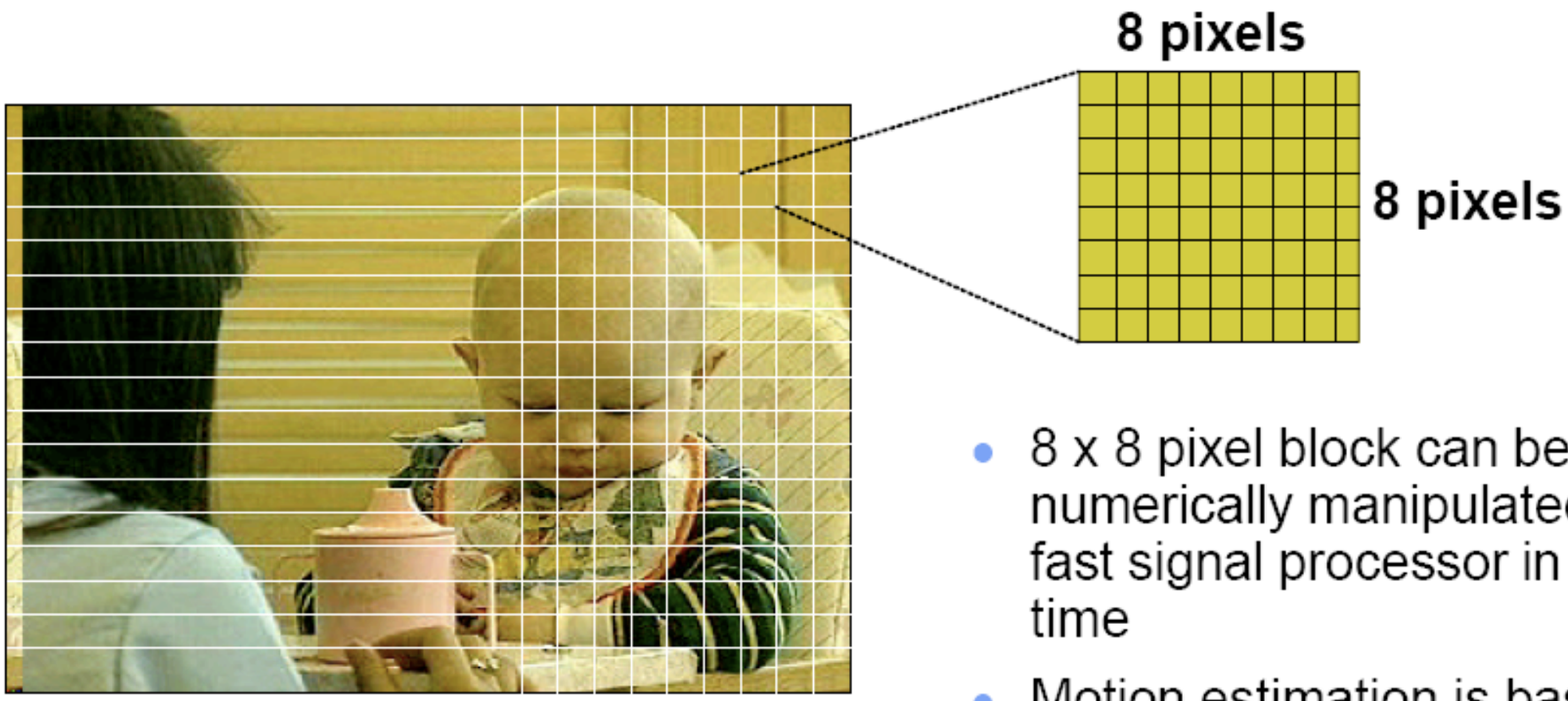
MPEG-2 Standard ISO/IEC 13818-2 (1994)

- Audio
  - compatible multichannel extension of MPEG-1 audio

- System
  - video, audio and data multiplexing defines tow presentations:
    - ***Program Stream*** for applications using near error free media
    - ***Transport Stream*** for more error prone channels

- Applications
  - satellite, cable, and terrestrial broadcasting,
  - digital networks, and
  - digital VCR

MPEG compression is based on 8 x 8 pixel block processing



- 8 x 8 pixel block can be numerically manipulated by fast signal processor in real time
- Motion estimation is based on comparing the blocks between series of pictures

MPEG: only compress moving parts

new picture

-

previous picture

=

difference

Encoder

Decoder

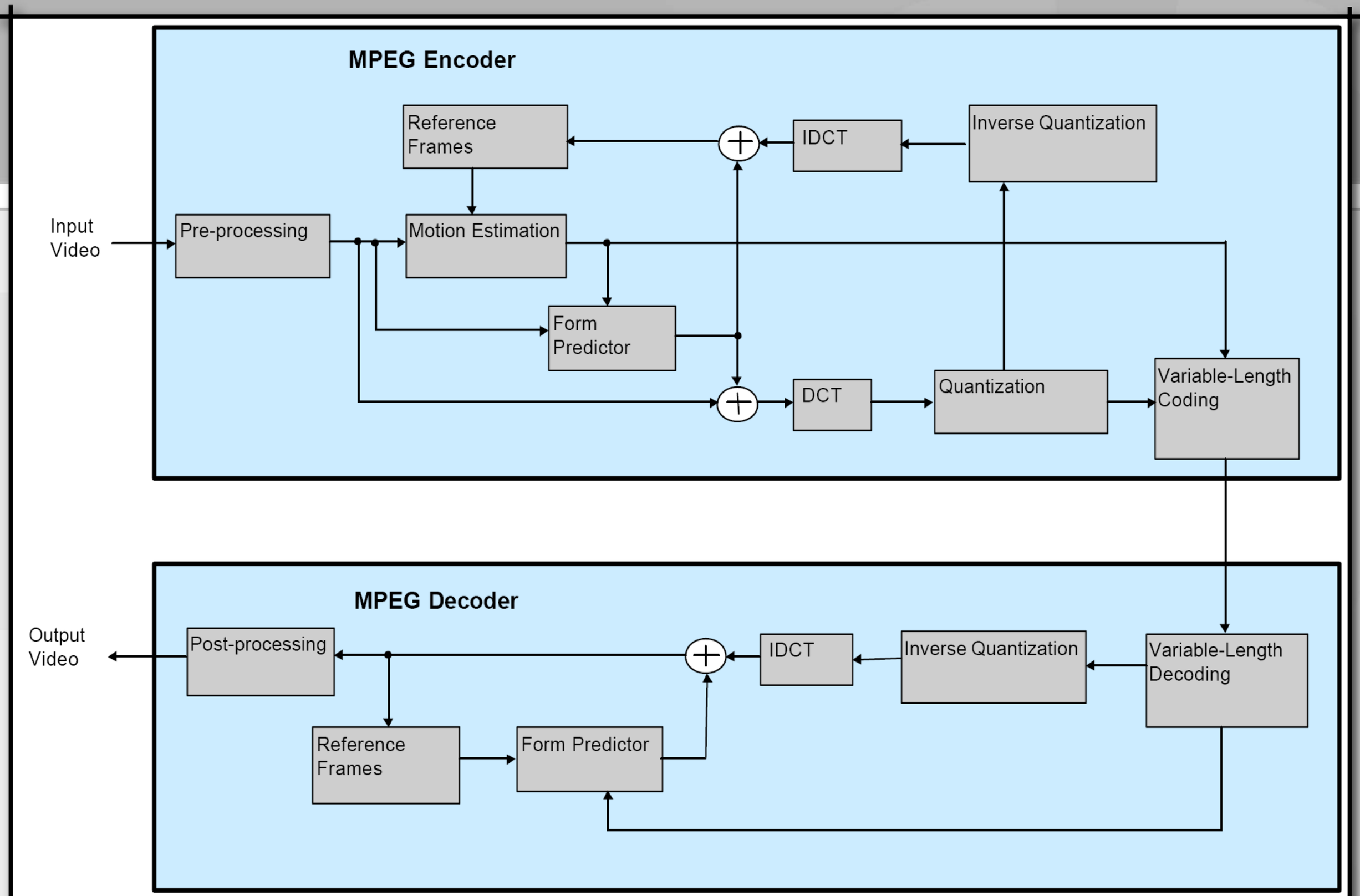
difference

+

previous picture

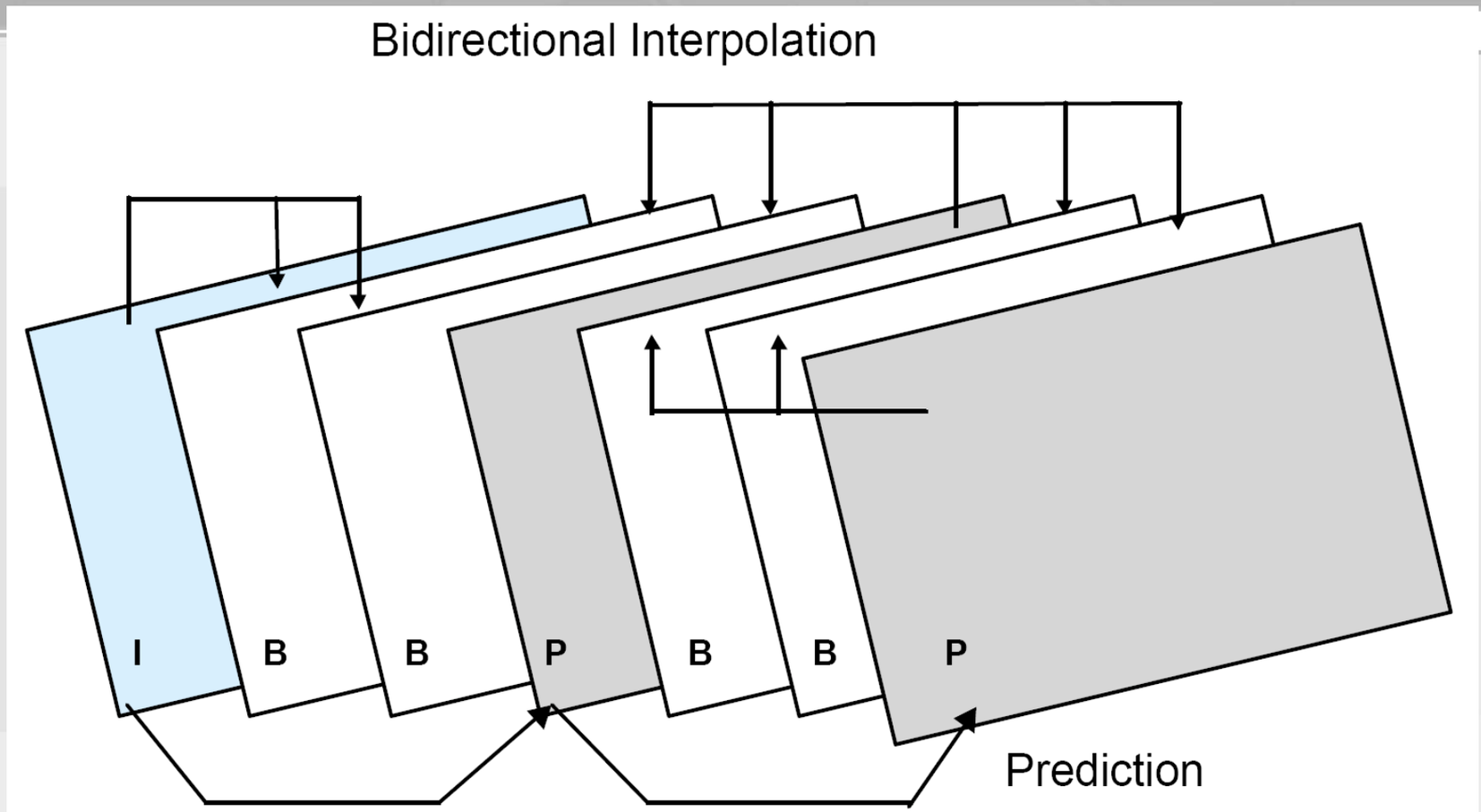
=

new picture


MPEG Encoder
Input Video
Pre-processing
Motion Estimation
Reference Frames
Form Predictor
IDCT
Inverse Quantization
DCT
Quantization
Variable-Length Coding
MPEG Decoder
Output Video
Post-processing
Reference Frames
Form Predictor
IDCT
Inverse Quantization
Variable-Length Decoding

MPEG: motion compression

I = Intra-Frame
B = Bi-directionally interpolated frame
P = Predicted frame



video signal: stream of picture, it is not necessary to send every picture

- Whole picture is needed only when all the content is changed!
- Several pictures has to be buffered to memory to make prediction forward and backward

MPEG: other issues

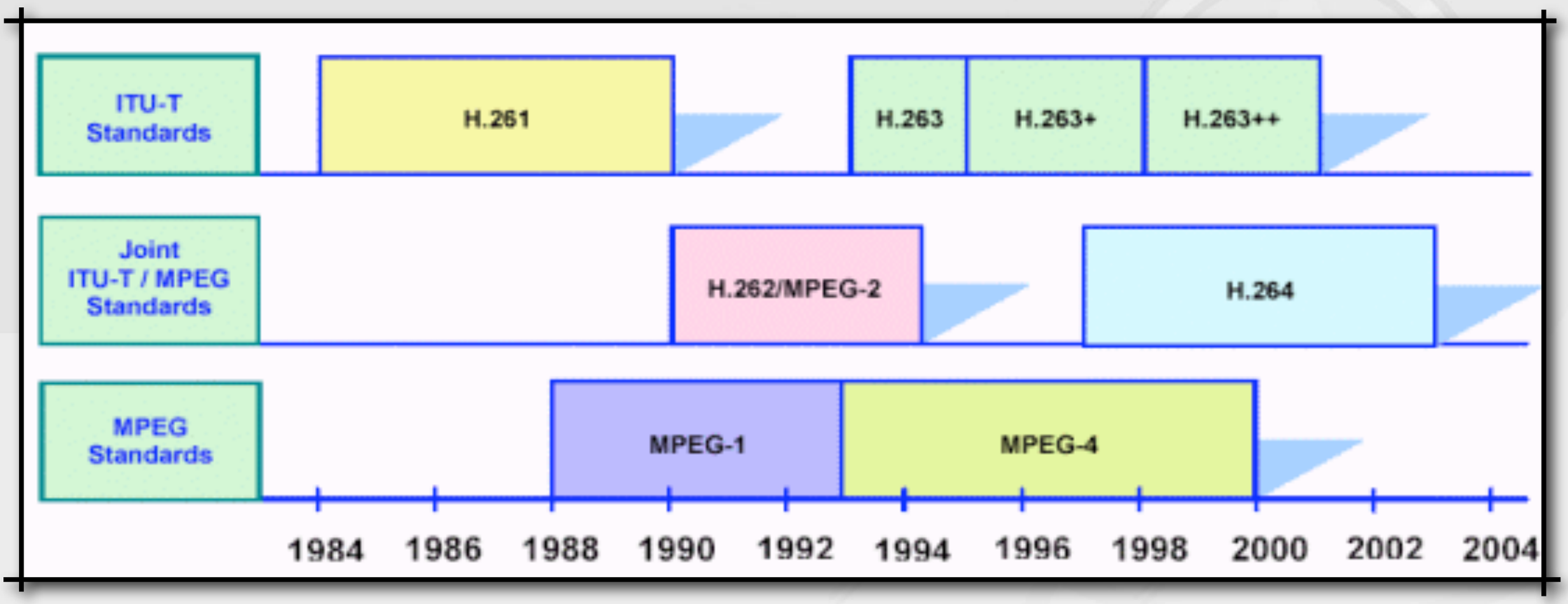
- Motion compensating
- Intra-frame transfer order

Overview of H.264

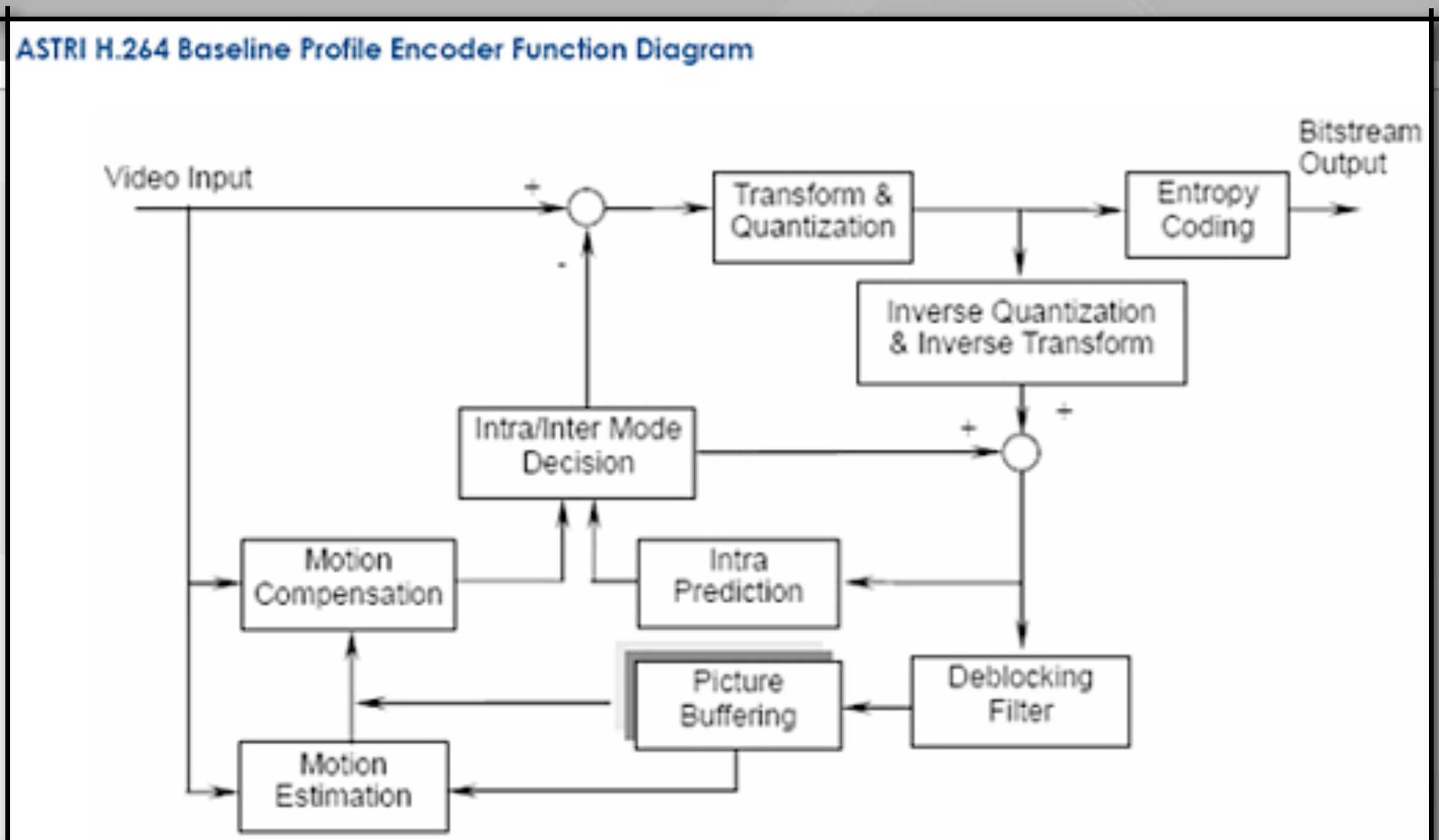
- JVT（Joint Video Team）
  - founded on December 2001, Pattaya Thailand.
  - video coding specialists from ITU-T and ISO, the two international standards organizations
  - goal: define a new video coding standards to achieve high compression rate, high image quality, good network adaptive coding frame.

- H.264: A new video compression standard
  - accepted by ITU-T
  - accepted by ISO
    - called AVC (Advanced Video Coding) standard
    - as the 10th part of MPEG-4

Major history of digital video standard



H.264 coding principle



H.264的主要技术特点

1. 4类DCT整数变换以及相应的量化方法

2. 7种宏块预测模式
   - 16×16, 16×8, 8×16, 8×8, 8×4, 4×8, 4×4
   - 运动估计和补偿更加精确

3. 多参考帧
4. 帧内预测
5. 改进的去块效应滤波器（Deblocking filter）
6. 增强的熵编码方法
   - UVLC（Universal VLC）、CAVLC（Context adaptive VLC）和CABAC

7. 1/4像素插值
8. 宏块级逐行、隔行自适应编码MBAFF

Advantages and shortages of H.264

- ☑ High compression rate
  - In the same image quality, H.264 can be compressed as size of
    - 36% of MPEG-2， 61% of MPEG-4 ， 51% of H.263
  - Low bit stream, high quality

- ☑ High error correctness rate
  - H.264 provides necessary tools to solve the error coding problem in unstable network environments

- ☑ Network adaptation
  - H.264 provides Network Adaptation Layer so as to make files of H.264c can be easily transferred in different network environments.

- ☐ High computation price
  - In the same image quality, H.264 is twice of MPEG-2 in computation complexity.

Applications of H.264

- H.264 standards added a NAL（Network Abstraction Layer）
  - to face the network connection and interface problem in the real applications.

- video communication
  - In real-time communication, POLYCOM、TANDBERG、VCON、SONY etc. claimed their own H.264 based TV-meeting products.

- digital TV broadcasting
  - MPEG has already finished defining the MPEG-2 compatible standard on H.264 stream coding content

- video storage-and-play-back
  - For High resolution DVD（HD DVD）application, H.264/MPEG-4 AVC solution.

Summary of video coding

- Resolution
- Coding rate
- Motion coding
- Transfer performance

浙江大学计算机学院
数字媒体与网络技术

2.4. HTML and XML

Overview of HTML

- Hypertext Markup Language
  - Developed by Tim Berners-Lee.
    - lightweight markup language vs. complex SGML.
  - Based on pure text format

- Rich abilities to display multimedia information.
  - Later added tags to support image and videos.

- HTML 3.2 => HTML 4.0 => HTML 5.0
  - Different browser has their own display effects.

Overview of all HTML elements

Reference: http://htmlhelp.com/reference/wilbur/overview.html

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<html>
<head>
   <title>Apple中国</title>
   <meta http-equiv="content-type" content="text/html;
charset=gb2312">
   ...
</head>

<body>
<!-- Tag for Activity Group: General, Activity: Apple China -
Homepage -->
...
</body>
</html>
```

http://www.apple.com.cn/

Overview of all HTML elements

Reference: http://htmlhelp.com/reference/wilbur/overview.html

Head →

Body →

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<html>
<head>
    <title>Apple中国</title>
    <meta http-equiv="content-type" content="text/html;
charset=gb2312">
    ...
</head>

<body>
<!-- Tag for Activity Group: General, Activity: Apple China -
Homepage -->
...
</body>
</html>
```

http://www.apple.com.cn/

Overview of HTML - Head elements

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html4/loose.dtd">
<html>
<head>
    <title>Apple中国</title>
    <meta http-equiv="content-type" content="text/html; charset=gb2312">
    ...
    <link rel="home" href="http://www.apple.com/">
    ...
    <script src="http://images.apple.com/global/scripts/lib/prototype.js" type="text/javascript" charset="utf-8">
    </script>
    ...

    <style type="text/css" media="all">
    ...
    #billboard { width: 1100px; margin: 0 auto 15px; overflow: hidden; position: relative; }
    #ticker { margin-bottom: 15px; }
    #homefooter { margin: 60px auto 50px; }
    ...
    </style>
</head>
```

- TITLE - Document title
- ISINDEX - Primitive search
- META - Meta-information
- LINK - Site structure
- BASE - Document location
- SCRIPT - Inline script
- STYLE - Style information

Overview of HTML - Body elements

```
<html>
<head> ... </head>
<body>
    <H1> Hello, world </H1>
    <P> Digital Asset management is cool! </P>

</body>
</html>
```

- **Block level elements**
  - –Headings: H1 => H6
  - –Lists: UL, OL, DIR, MENU, LI, DL, DT, DD
  - –Text Containers: P, PRE, BLOCKQUOTE, ADDRESS
  - –others: DIV, CENTER, FORM, HR, TABLE

Overview of HTML - Body elements

```
<html>
<head> ... </head>
<body>
    <H1> Text-level elements </H1>
    <A href="http://www.google.com"> GOOGLE  <IMG src=" ... ">  </A>

</body>
</html>
```

- **Text-level elements**
  - –Logical markup: EM ...
  - –Special markup: A, IMG, APPLET ...
  - –Physical markup: B, ...
  - –Forms: INPUT ...
  - –Tables: CAPTION, TR, TH, TD

Overview of XML

- Extensible Markup Language
  - Aim at data searching

  –Similar to HTML
    - More restrict grammar checking
    - User defined tags to describe data structure
    - Flexible data displaying schemes
    - Cross-platform, language and application independent
    - DTD and XML Schema.

- http://www.brics.dk/~amoeller/XML/overview.html

HTML v.s. XML

```
<h1>Rhubarb Cobbler</h1>
<h2>Maggie.Herrick@bbs.mhv.net</h2>
<h3>Wed, 14 Jun 95</h3>

Rhubarb Cobbler made with bananas as the main sweetener.
It was delicious.  Basicly it was
  <table>
  <tr><td> 2 1/2 cups <td> diced rhubarb
  <tr><td> 2 tablespoons <td> sugar
  <tr><td> 2 <td> fairly ripe bananas
  <tr><td> 1/4 teaspoon <td> cinnamon
  <tr><td> dash of <td> nutmeg
  </table>
Combine all and use as cobbler, pie, or crisp.
Related recipes: <a href="#GardenQuiche">Garden Quiche</a>
```

```
<recipe id="117" category="dessert">
  <title>Rhubarb Cobbler</title>
  <author><email>Maggie.Herrick@bbs.mhv.net</email></author>
  <date>Wed, 14 Jun 95</date>

  <description>
    Rhubarb Cobbler made with bananas as the main sweetener.
    It was delicious.
  </description>

  <ingredients>
    <item><amount>2 1/2 cups</amount><type>diced rhubarb</type></item>
    <item><amount>2 tablespoons</amount><type>sugar</type></item>
    <item><amount>2</amount><type>fairly ripe bananas</type></item>
    <item><amount>1/4 teaspoon</amount><type>cinnamon</type></item>
    <item><amount>dash of</amount><type>nutmeg</type></item>
  </ingredients>

  <preparation>
    Combine all and use as cobbler, pie, or crisp.
  </preparation>

  <related url="#GardenQuiche">Garden Quiche</related>
</recipe>
```

A conceptual view of XML

```
...<foo attr="val" ...>...</foo>...
```

a matching element end tag

the contents of the element

an attribute with name attr and value val, values enclosed by ' or "

an element start tag with name foo

```
...<foo attr="val" ... />...
```

XML documents as text with markup

A conceptual view of XML

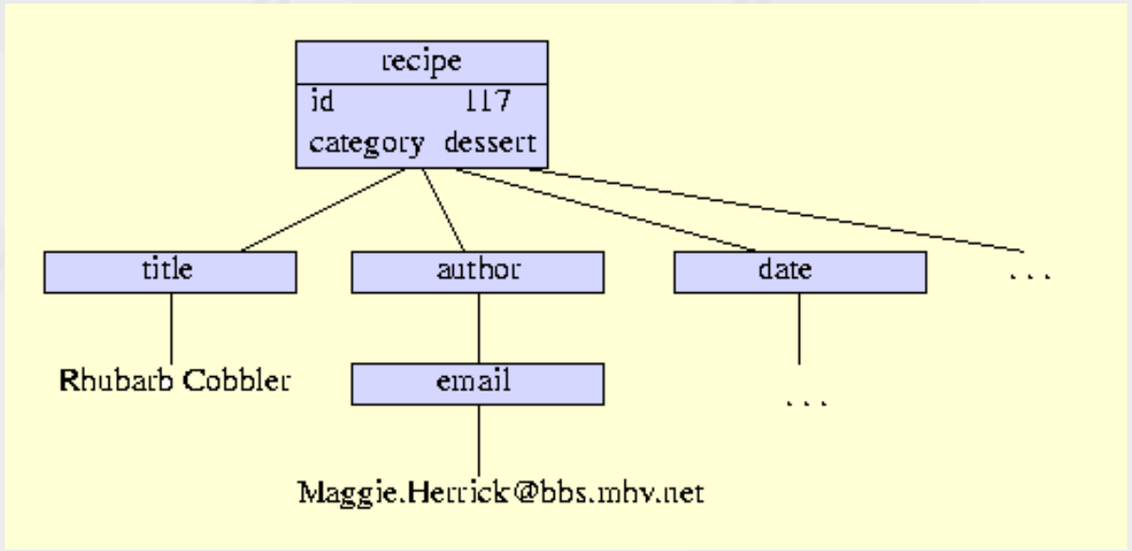
- An XML document is a (Unicode) text with markup tags and other meta-information.
- An XML document must be well-formed:
  - start and end tags must match
  - element tags must be properly nested
  - + some more subtle syntactical requirements

- XML is case sensitive!
- Special characters can be escaped using Unicode character references:
  - < and < both yield <

A conceptual view of XML

- An XML document is an ordered, labeled tree:
  - character data leaf nodes contain the actual data (text strings)
    - usually, character data nodes must be non-empty and non-adjacent to other character data nodes
  - elements nodes, are each labeled with
    - a name (often called the element type), and
    - a set of attributes, each consisting of a name and a value,
    - and these nodes can have child nodes

XML documents as labeled trees



A conceptual view of XML

- XML trees may contain other kinds of leaf nodes:
  - processing instructions - annotations for various processors
  - comments - as in programming languages
  - document type declaration

XML documents as labeled trees

- The XML vision offers:
  - common extensions to the core XML specification
    - a namespace mechanism, document inclusion, etc.
  - schemas
    - grammars to define classes of documents
  - linking between documents
    - a generalization of HTML anchors and links
  - addressing parts of read-only documents
    - flexible and robust pointers into documents
  - transformation
    - conversion from one document class to another
  - querying
    extraction of information, generalizing relational databases

To use XML

- Define your XML language
  - use XML Schema to define its syntax
- Exploit the generic XML tools
  - XSLT and XQuery processors
- As a generic protocols, and the generic programming frameworks
  - DOM or SAX to build application tools

Summary: HTML and XML

- Both of them are useful today for different applications

浙江大学计算机学院
数字媒体与网络技术

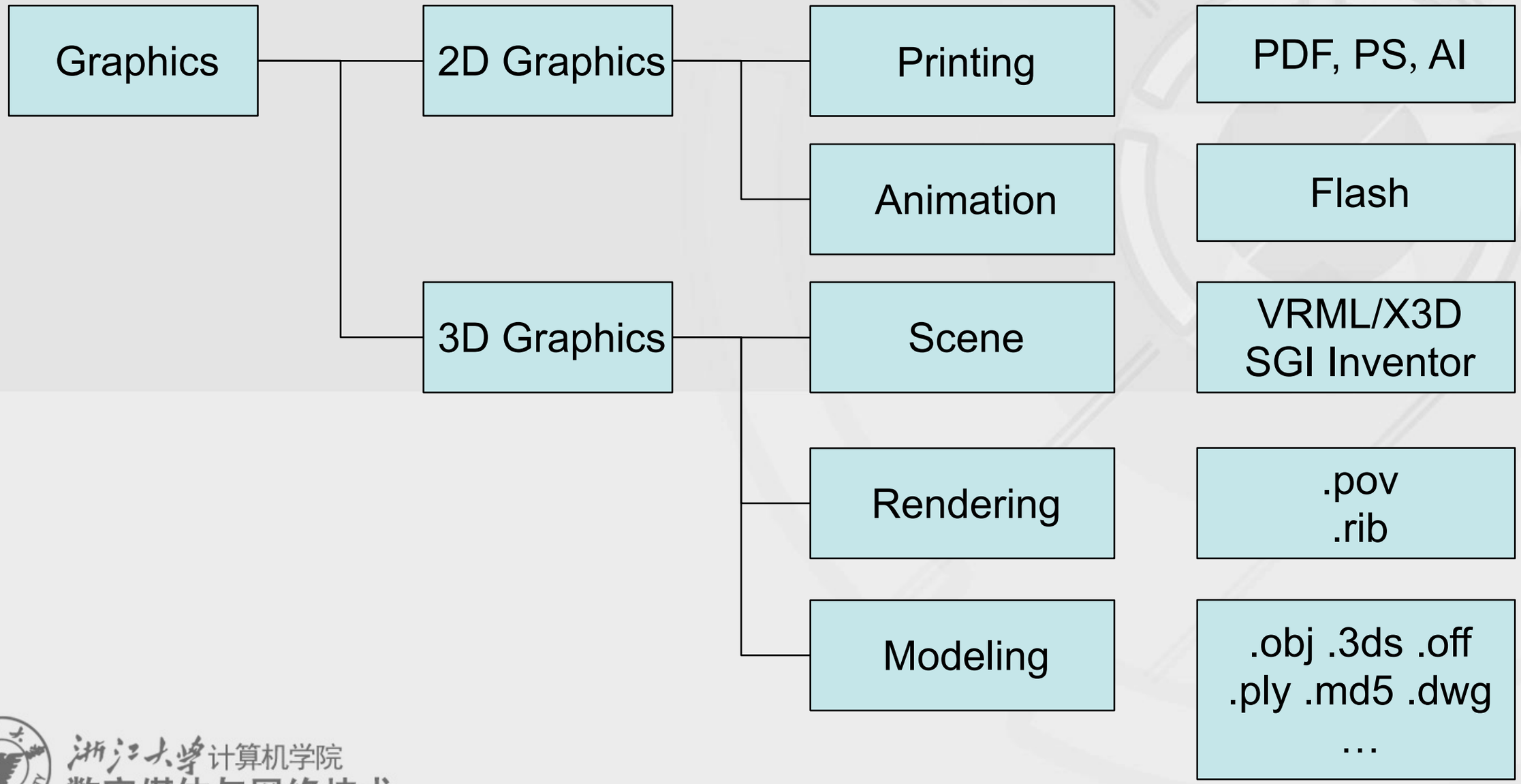
2.5. Graphics formats

Graphics ≠ Images

- Representation ability
  - Graphics are usually described in vectors which can provide arbitrary precision
  - Images are usually sampled in fragments/pixels which can only provide limited precision

- Application area
  - Graphics are mainly applied in CAD, model design, computer animation, system simulation and printing.
  - Images are mainly used for photo display and image processing etc.

Classification of different graphics formats



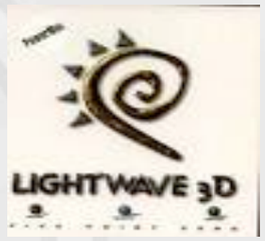
Overview of FLASH

Elements of 3D graphics format

- Global scene description
  - Parameters of light and camera, other system configurations
- Geometric model description
  - Curves and surfaces
    - Line, plane, quadratic surface, spline …
  - Mesh surfaces = vertex coordinates + topology connectivity
  - Texture coordinates, normals
- Material description
  - Reflectance model, texture image
- Animation description
  - Skeleton model …

Main problems for 3D graphics format

- CAD and computer animation software
  - Different application area
  - Different system design principles
  - Different types of geometric representation combinations



- Mainstream commercial software employ different types of 3D graphics model.
  - It is hard to obtain a uniform graphics format.
  - Data exchange and sharing become key issues for 3D designing system.

Overview of X3D

- X3D [Extensible 3D] is an international standard of 3D graphics. It defines how to integrate and access interactive 3D content in a multimedia environment.

- The former of X3D is VRML which is established on 1998 as a network graphics ISO standard (ISO/IEC14772).

- X3D decompose scene descriptions of VRML97 into components. Therefore it is very convenient to extend original VRML functions by adding new components.

New 3D graphics standard-X3D

- Ten years from VRML to X3D

1994.10　通过VRML1.0 三维文件格式
1996.7　公布VRML2.0 草案加入交互特性
1998.1　通过VRML97国际标准
1998.11　改名为Web3D联盟，推荐结合
1999.2　启动X3D
1999 - 2002 实现了 gzip、Universal-Media-Libraries、GeoVRML、DIS-Java-VRML、H-Anim、EAI
2002.4　VRML标准修订，正式加入UTF-8、EAI、GeoVRML、NURBS 曲面特性
2002.7　X3D 宣布草案
2002.12　X3D 进入ISO审议
2003.2　X3D 编码规格进入ISO审议
2003.3　X3D 语言结合标准进入ISO的最后审议阶段
2004　通过 X3D ISO 国际标准

3D mesh surface compression

- Terrain data can be compressed by JPEG related methods
- MPEG-4 defines a compression method:
  - Compress topological connectivities: relationships among vertices
  - Compress geometric position information: vertex positions, normal vectors, texture coordinates …
  - Compress texture images ...

Homework

- Build a simple image browser that can convert different types of images.