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

2010-09-14

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

浙江大学计算机学院
数字媒体与网络技术

2.1 Image format and coding methods

Common image formats

- General types:
  - –GIF
  - –JPEG
  - –PNG
  - –TIFF
  - –TGA
- Raw data:
  - –RAW
  - –DNG
- Platform spec.:
  - ➡BMP (Win)
  - ➡PAINT&PICT (Mac)
  - ➡PPM (X-Win)
- Vector data:
  - ➡WMF (Win)
  - ➡PS and PDF

浙江大学计算机学院
数字媒体与网络技术

Common image formats

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

LZW and lossless compression

- 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

Lossless image compression methods

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

Lossy image compression methods

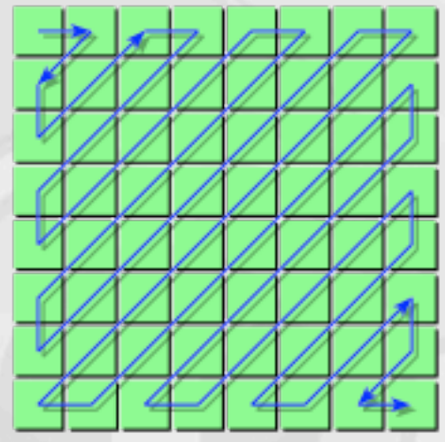
- Quantization
- Transform coding
  - –Discrete Cosine Transform => JPEG
  - –Discrete Wavelet Transform => JPEG 2000
  - –...

Image compression standards

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

浙江大学计算机学院
数字媒体与网络技术

JPEG compression: main idea



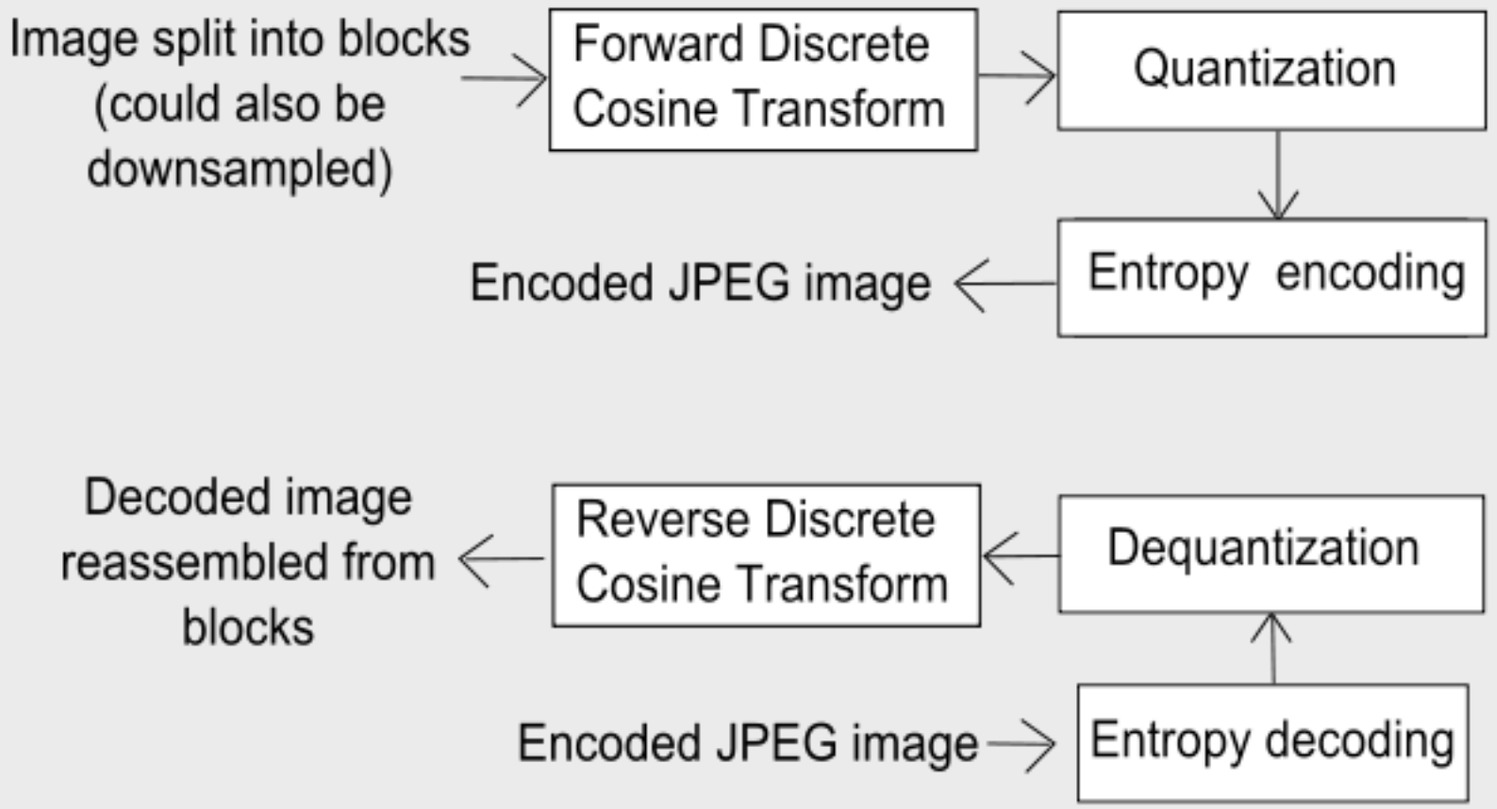
RGB Image → YCbCr Color Model → 8x8 image blocks → DCT → Quantization → frequency-domain representation

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

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

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

JPEG compression: implementation



Compression Rate

bpp: bit per pixel

Upper-left: The original image.
Upper-right: Decoded at 0.5 bpp (PSNR: 35.32 dB).
Lower-left: 1.0 bpp (PSNR: 38.73 dB).
Lower-right: 1.5 bpp (PSNR: 41.62 dB).

reference: http://cobweb.ecn.purdue.edu/~ace/color-wavelet/cwavelet.html

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

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)

DNG: Digital Negative

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

Summary – Essential factors of image storage

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

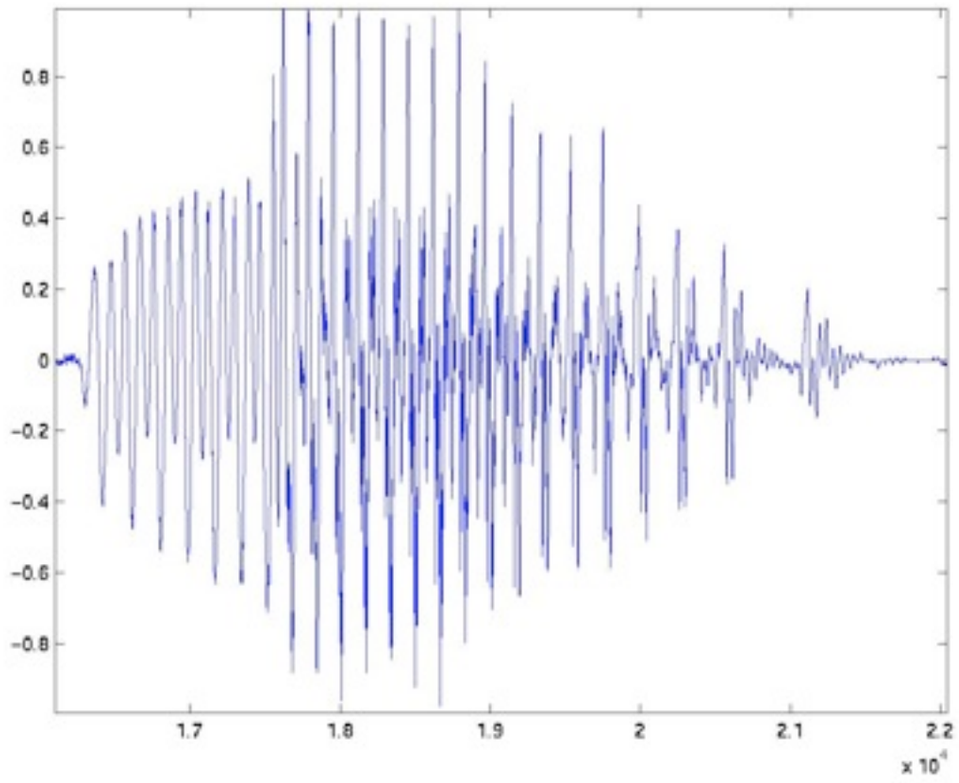
Image converting tools

- 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

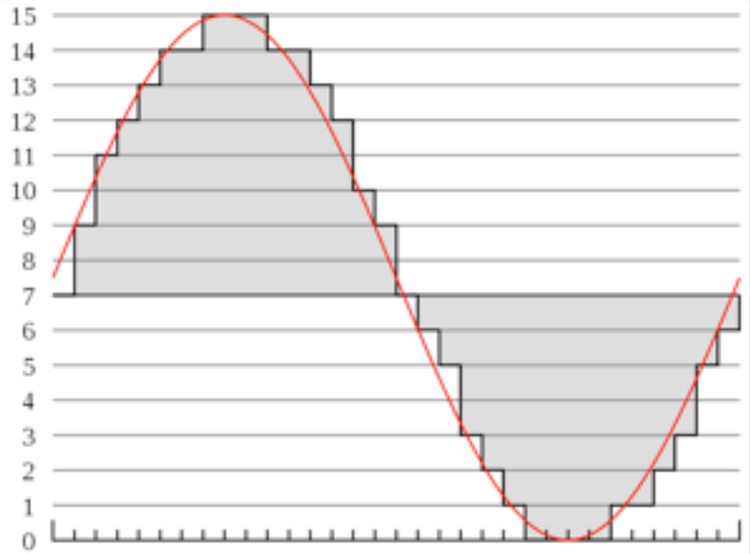
- There is no absolute lossless coding schemes!
  - According to the definitions of bit rate and quantization rate, audio coding can only approximate to the natural sound signal as much as possible.
  - Comparing with natural signal, all coding schemes are lossy.

- Related lossless 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 lossless coding scheme. However, 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.



Sampling and quantization of a signal (red) for 4-bit PCM

- Audio CD mainly leverage the PCM coding scheme.
  - One 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.

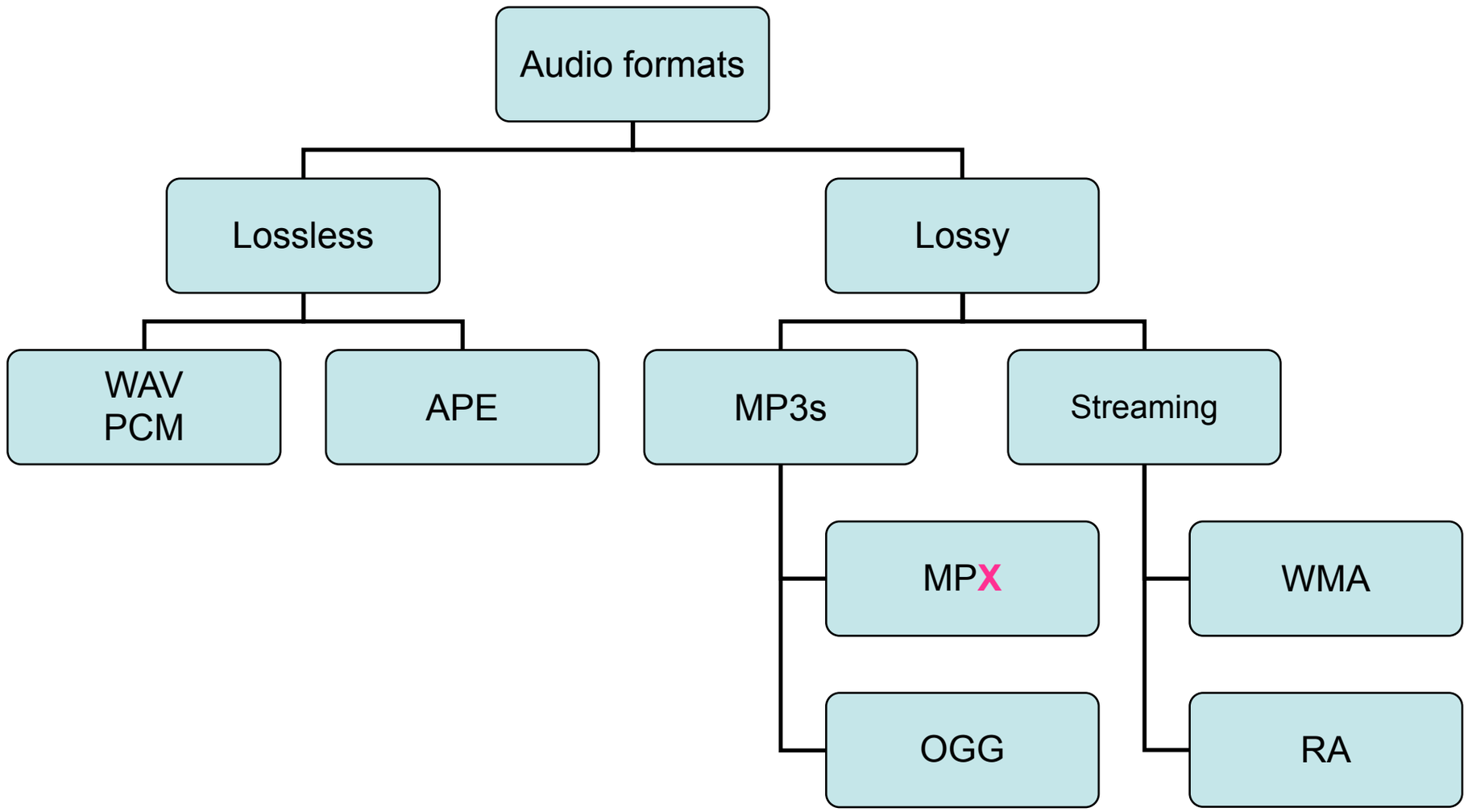
The streaming feature of audio

- 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 RIFF (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.

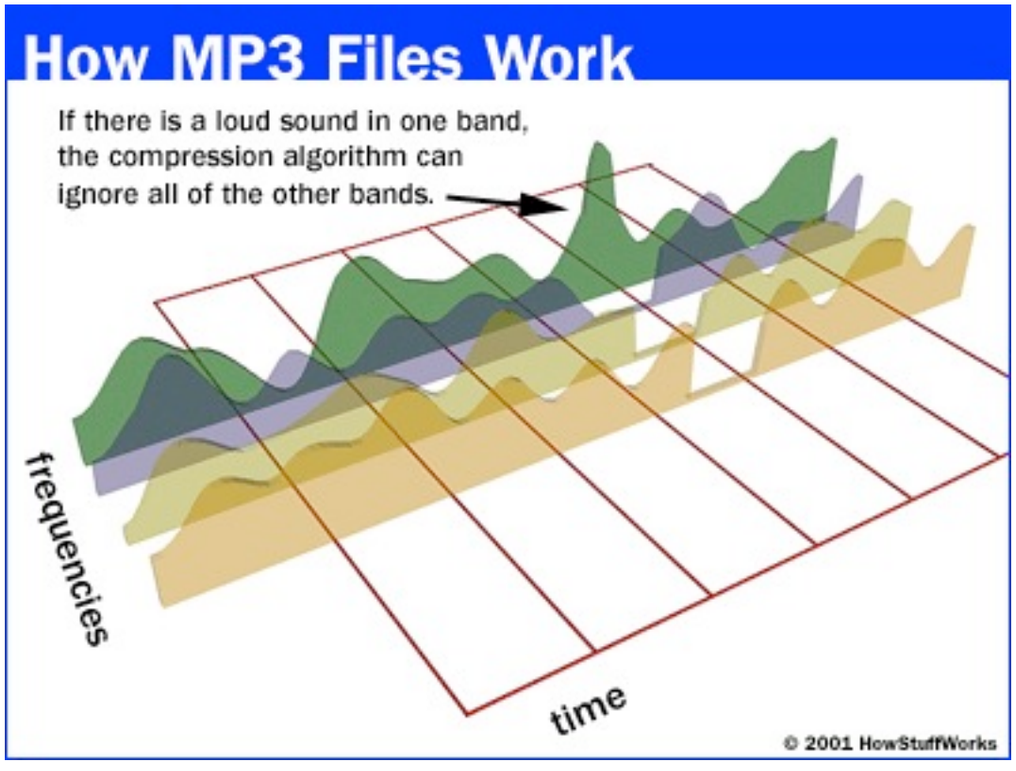
- RA mainly focus on network media market
  - Highlight: RA can alter its own coding bit rate due to the network width but keep the sound quality as much as possible.
  - RA can support many types of audio coding schemes, e.g., ATRAC3.
  - Beside the function of download-while-play, RA can also hide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

MP3

- From the MPEG-3 standard
- Most popular audio file format
- Special compression method for sound



perceptual noise shaping

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

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

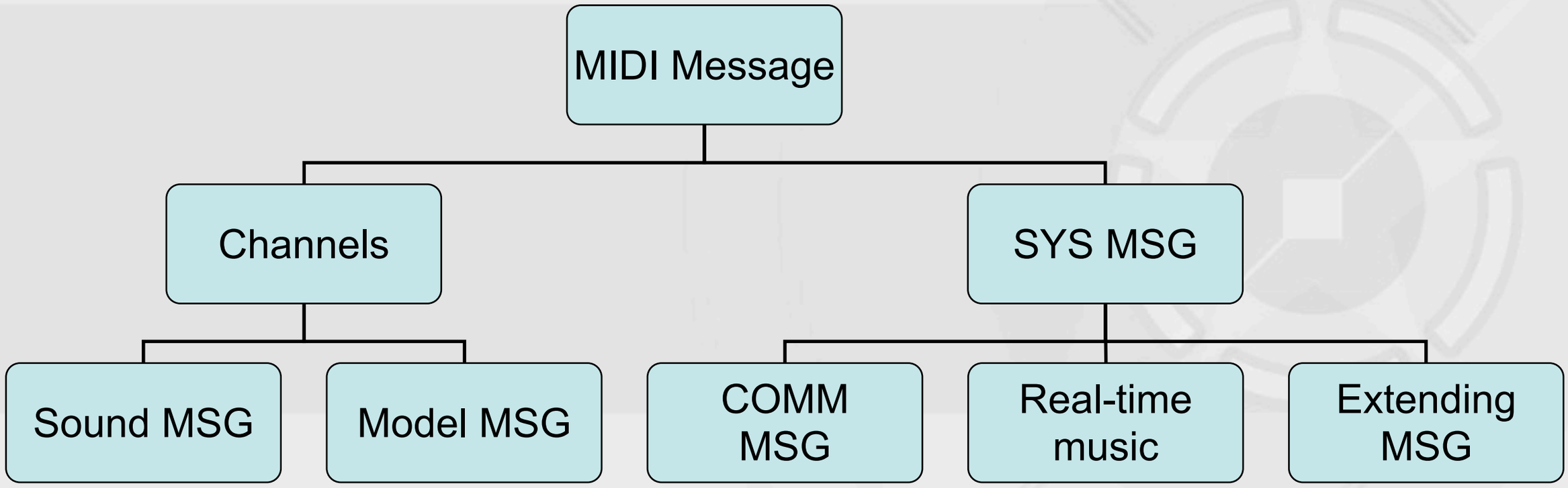
- MIDI is type of description language.
  - Different directly record digitalized sound signal
  - Only record ‘events’ that how instruments make sound.
  - Small storage size.

- 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

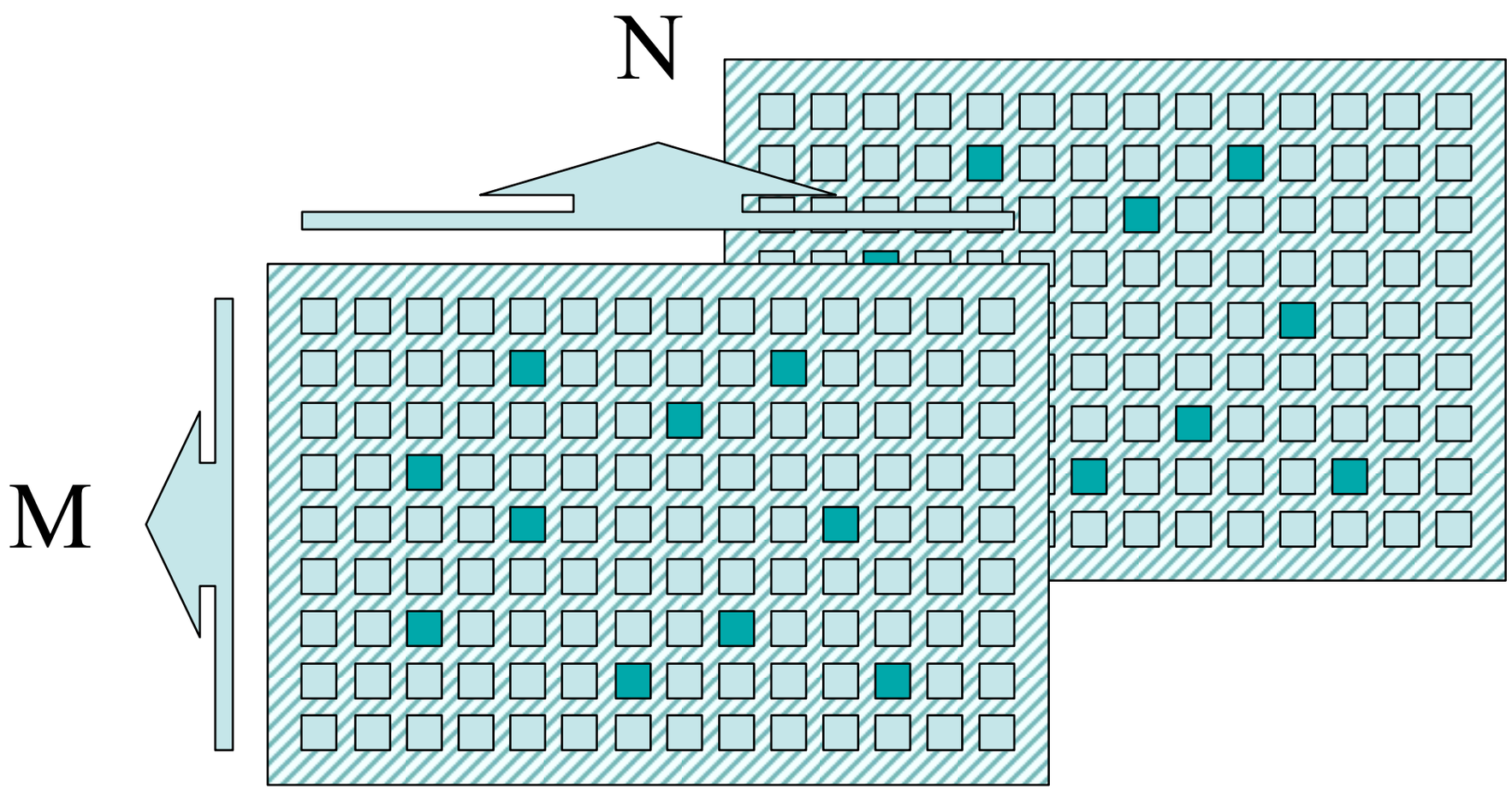
- MID
  - –General MIDI
- SMF
  - –Standard MIDI File

浙江大学计算机学院
数字媒体与网络技术

2.3. Video formats and coding methods

Representations of video

- Sequence of images ?!?
  - –Can be viewed as a 3-dimensional 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* (类) and *levels* （级）
    - 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

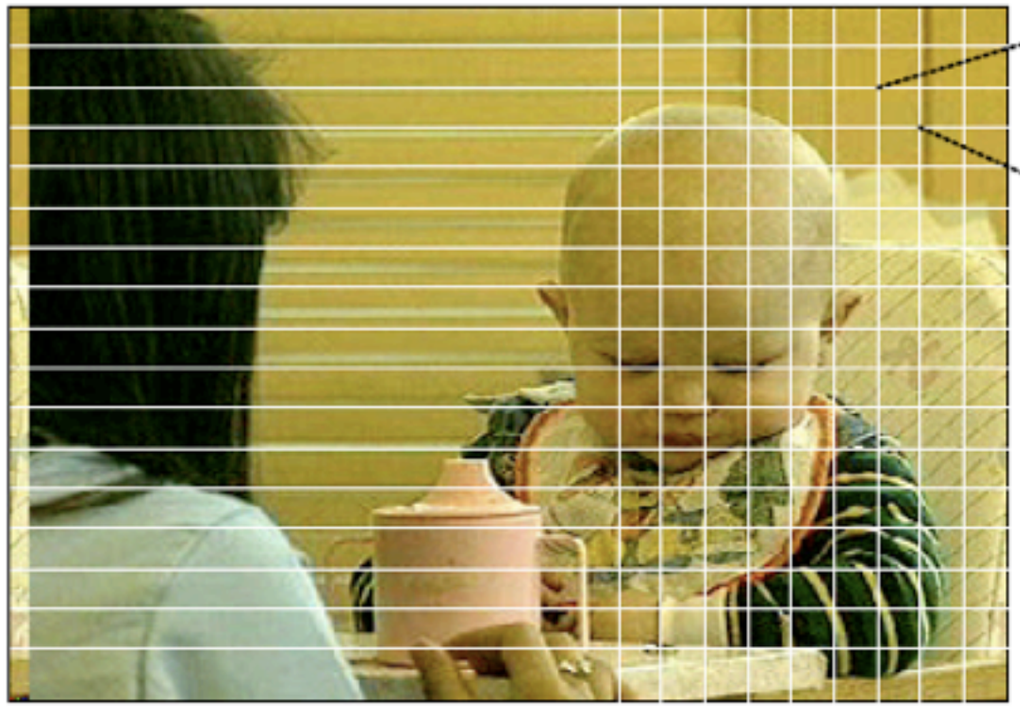
MPEG-2 Standard ISO/IEC 13818-2 (1994)

- 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

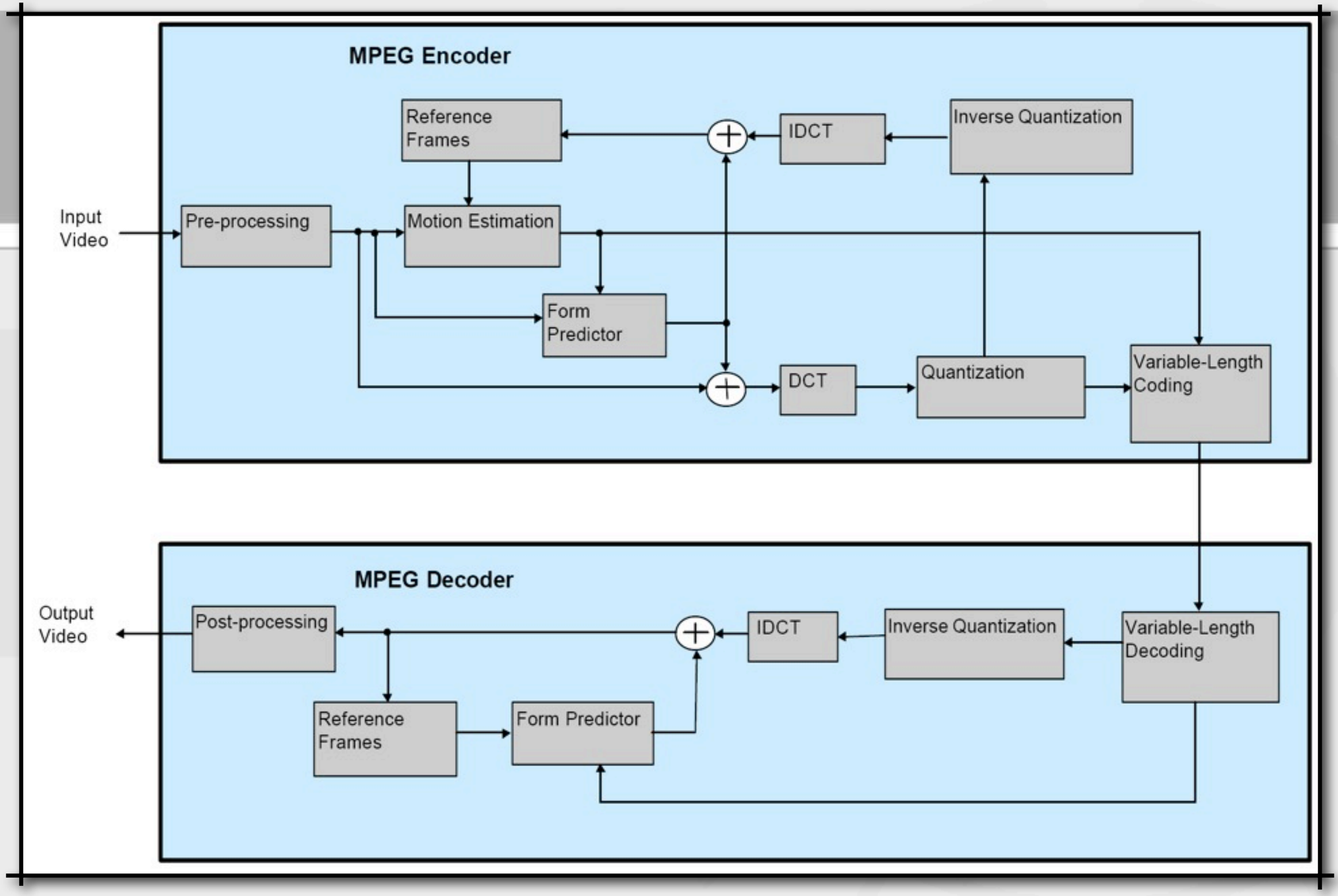
difference

+

previous picture

=

new picture

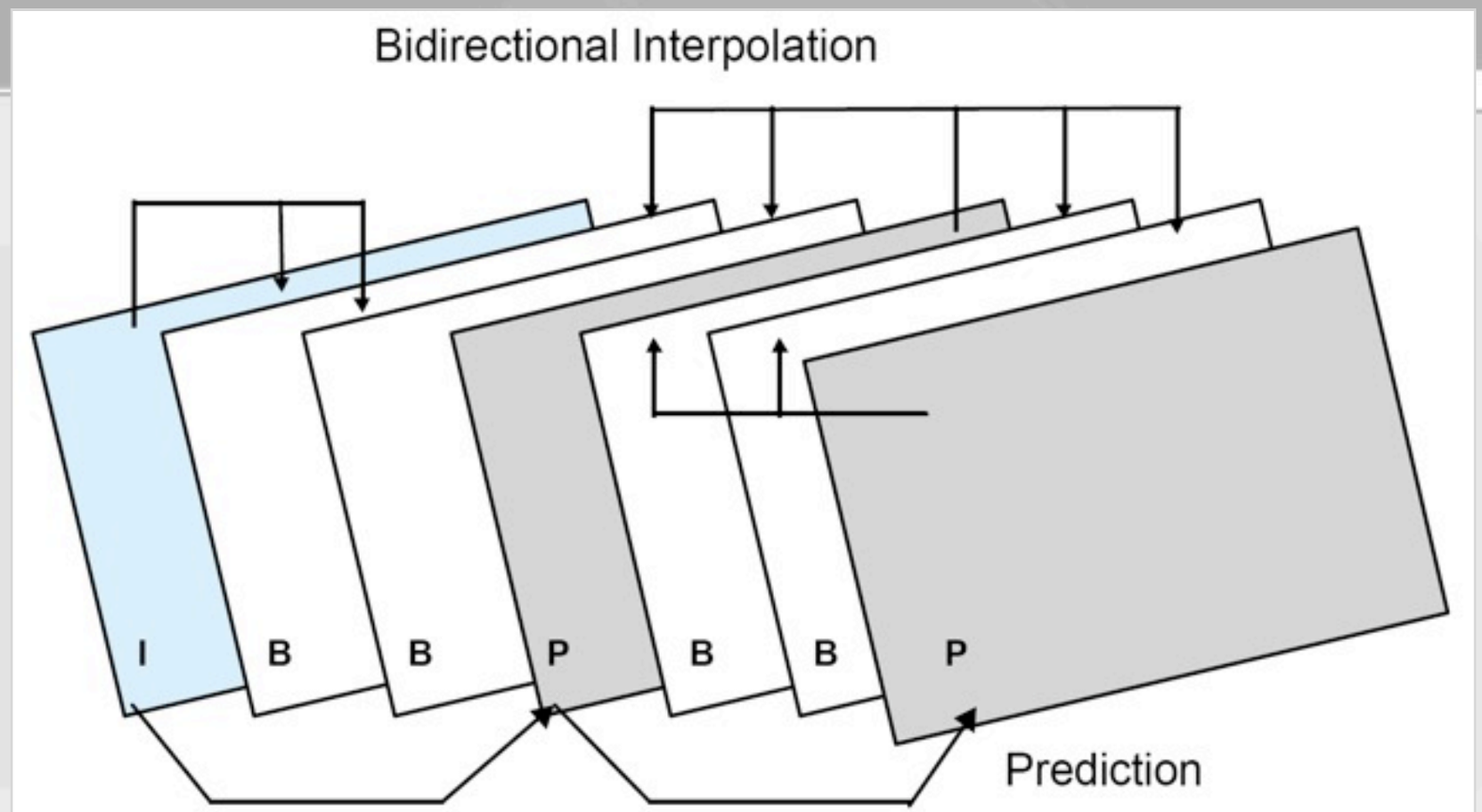


浙江大学计算机学院
数字媒体与网络技术

MPEG encoding and decoding

MPEG: motion compression

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



Typical sequence: IBBPBBPBBI

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

Color video coding

- 抽样和编码整个模拟（彩色）视频信号
  - –例如，复合编码

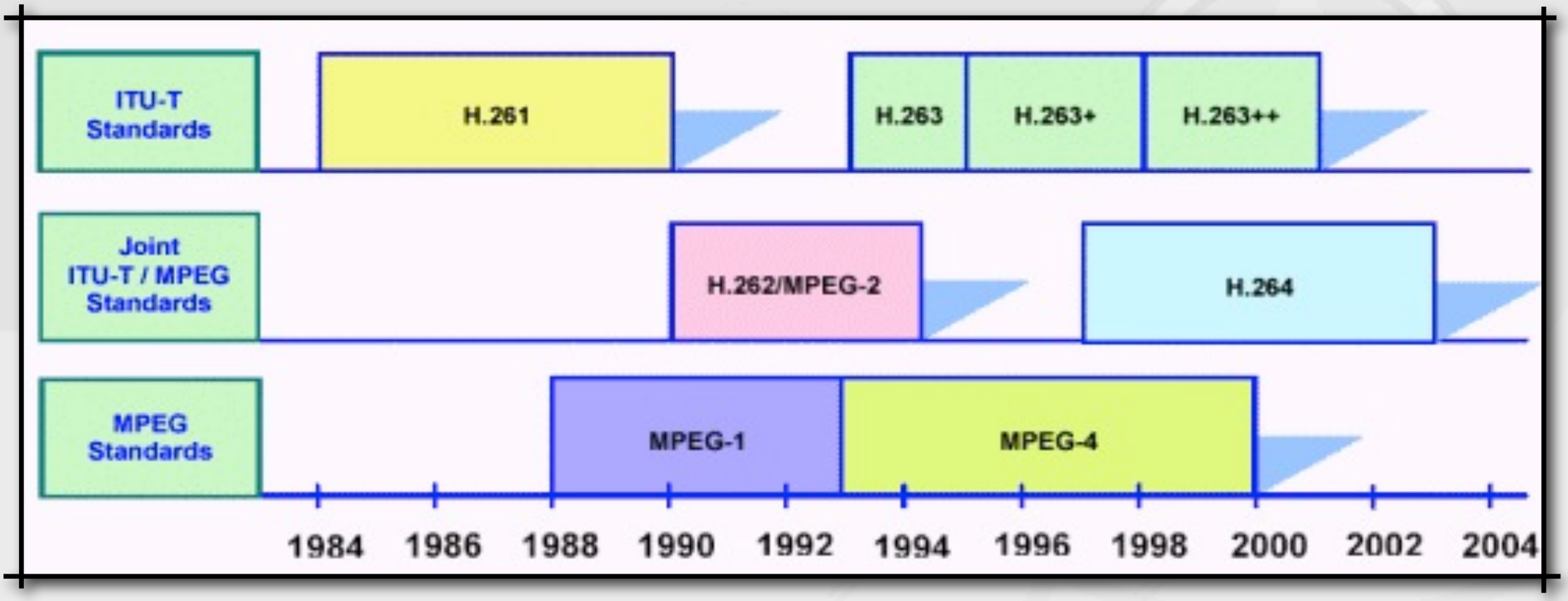
- 对亮度和色度分别编码
  - –例如，组件编码
  - –亮度比色度更重要，可根据应用场合采用4:2:2，4:2:0, 4:4:4等不同的编码比率

Overview of H.264

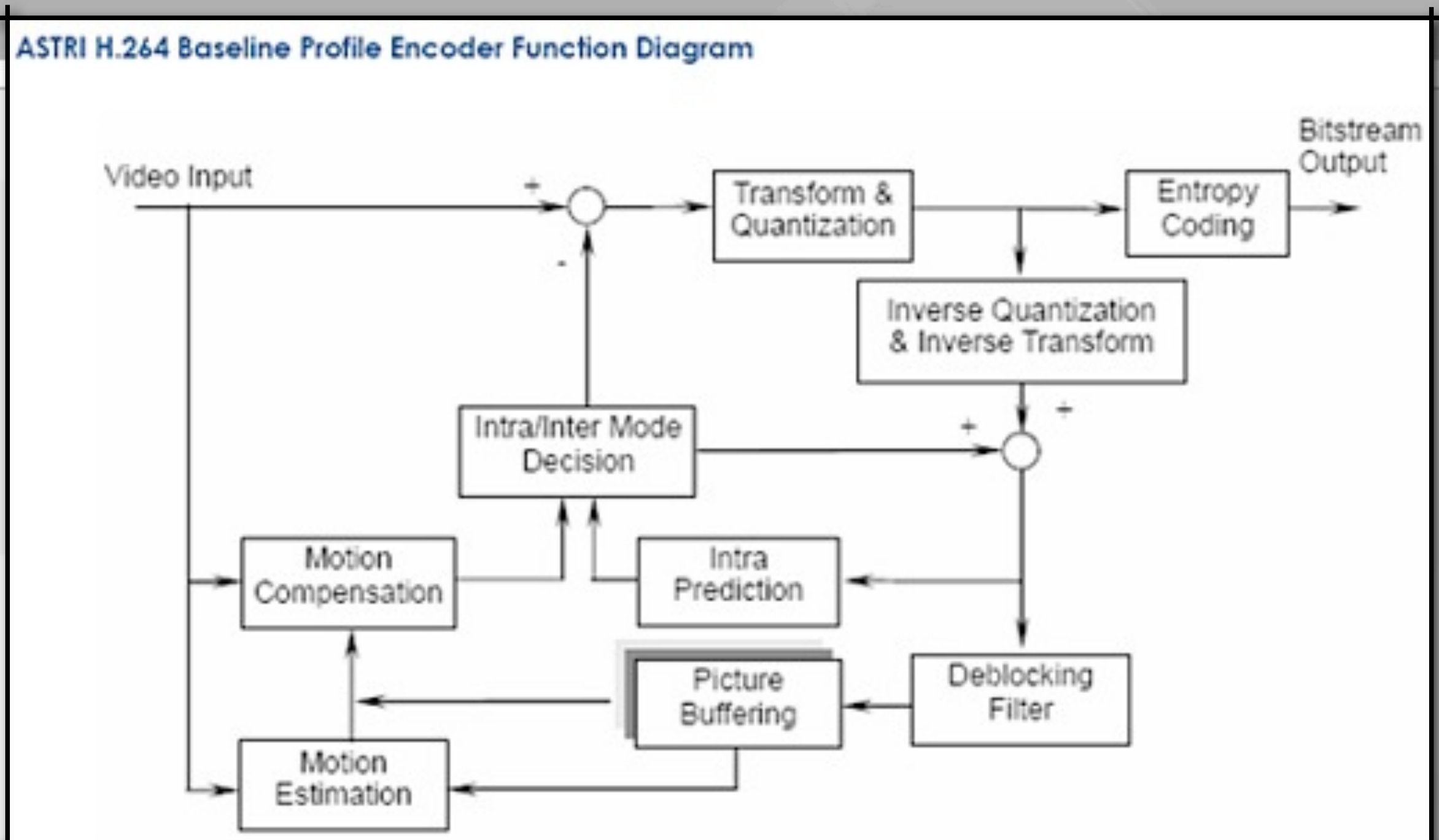
- JVT（Joint Video Team）
  - founded on December 2001, Pattaya Thailand.
  - video coding specialists from ITU-T and ISO, the two international standards organizations
  - goal: define a new video coding standards to achieve high compression rate, high image quality, good network adaptive coding frame.

- 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.

Video format Conversion

- Commercial tools
- Open source / Free tools
  - –http://handbrake.fr/

Summary of video coding

- Resolution
- Coding rate
- Motion coding
- Transfer performance