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

2011-09-16

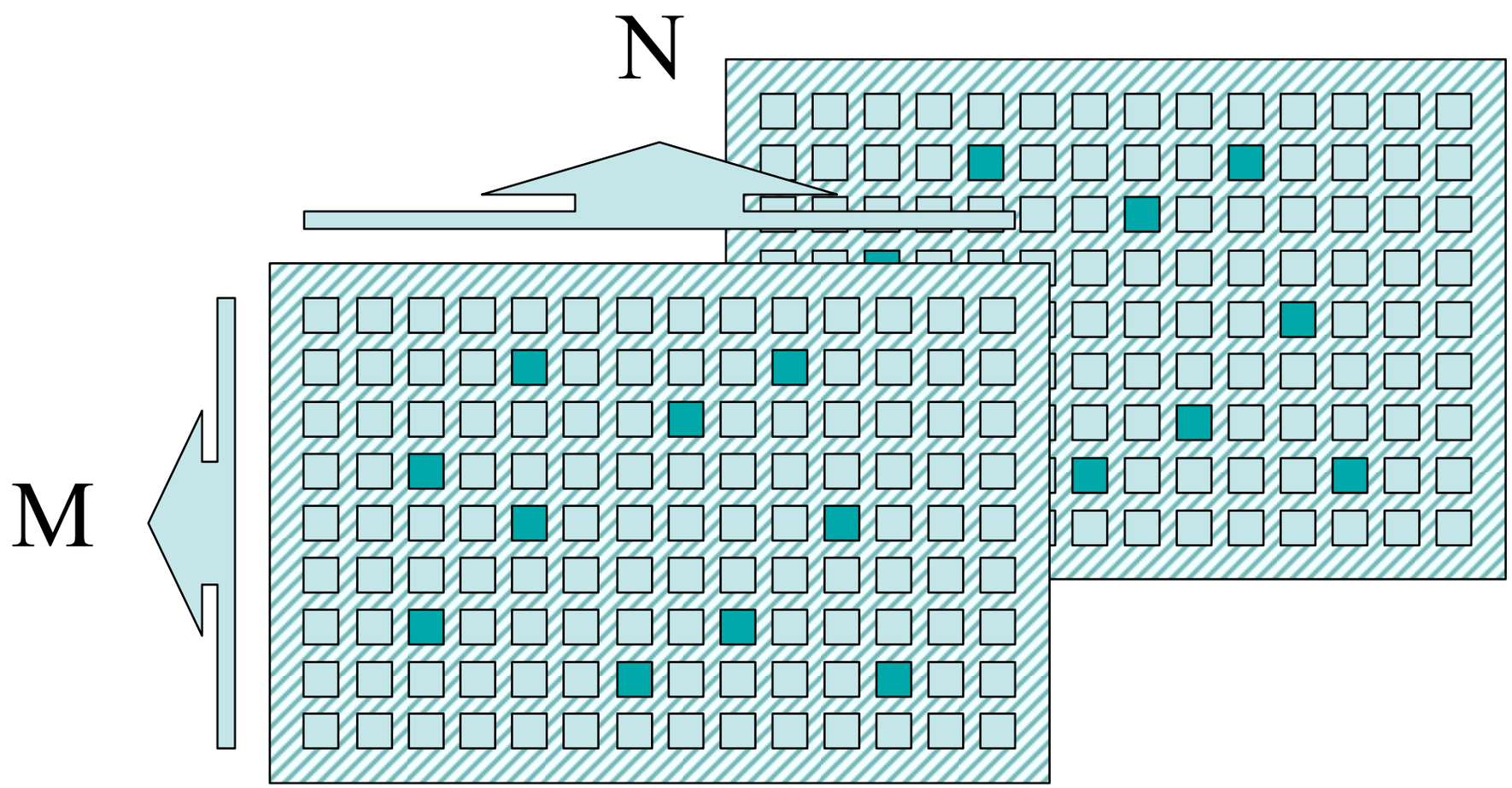
Outline

- 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

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/

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

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-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

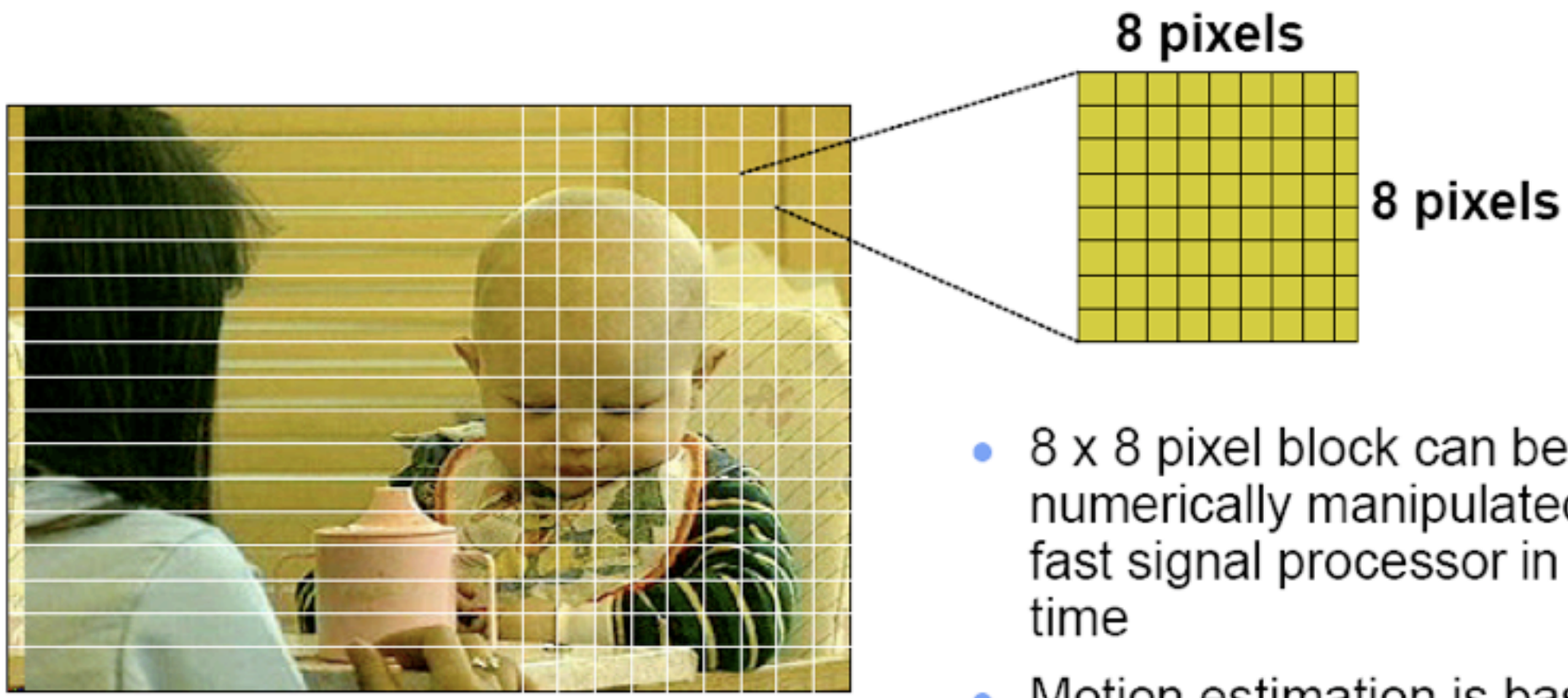
  - MPEG-2 consists of *profiles* (类) and *levels* （级）
    - Main Profile, Main Level (MP@ML)
      - 720x480 resolution video at 30 frames/sec
      - < 15 Mbit/sec (typical ~4 Mbit/sec)
      - for NTSC video

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

MPEG-1 v.s. MPEG-2

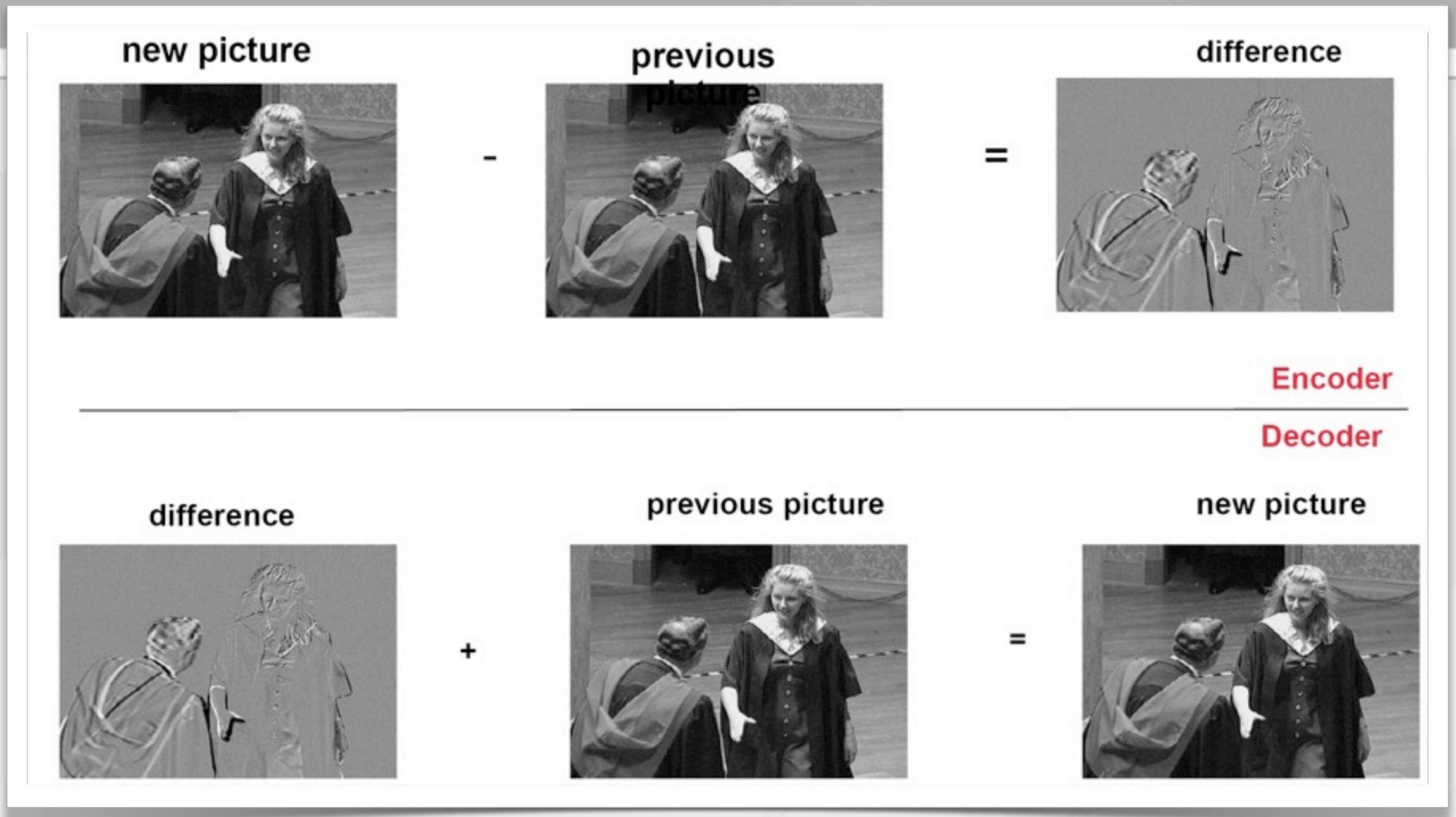
- MPEG-1 Apps ~
  - CD-I, digital multimedia,
  - video database, e.g. video-on-demand
- MPEG-2 Apps ~
  - satellite, cable, and terrestrial broadcasting,
  - digital networks, and
  - digital VCR

MPEG compression is based on 8 x 8 pixel block processing



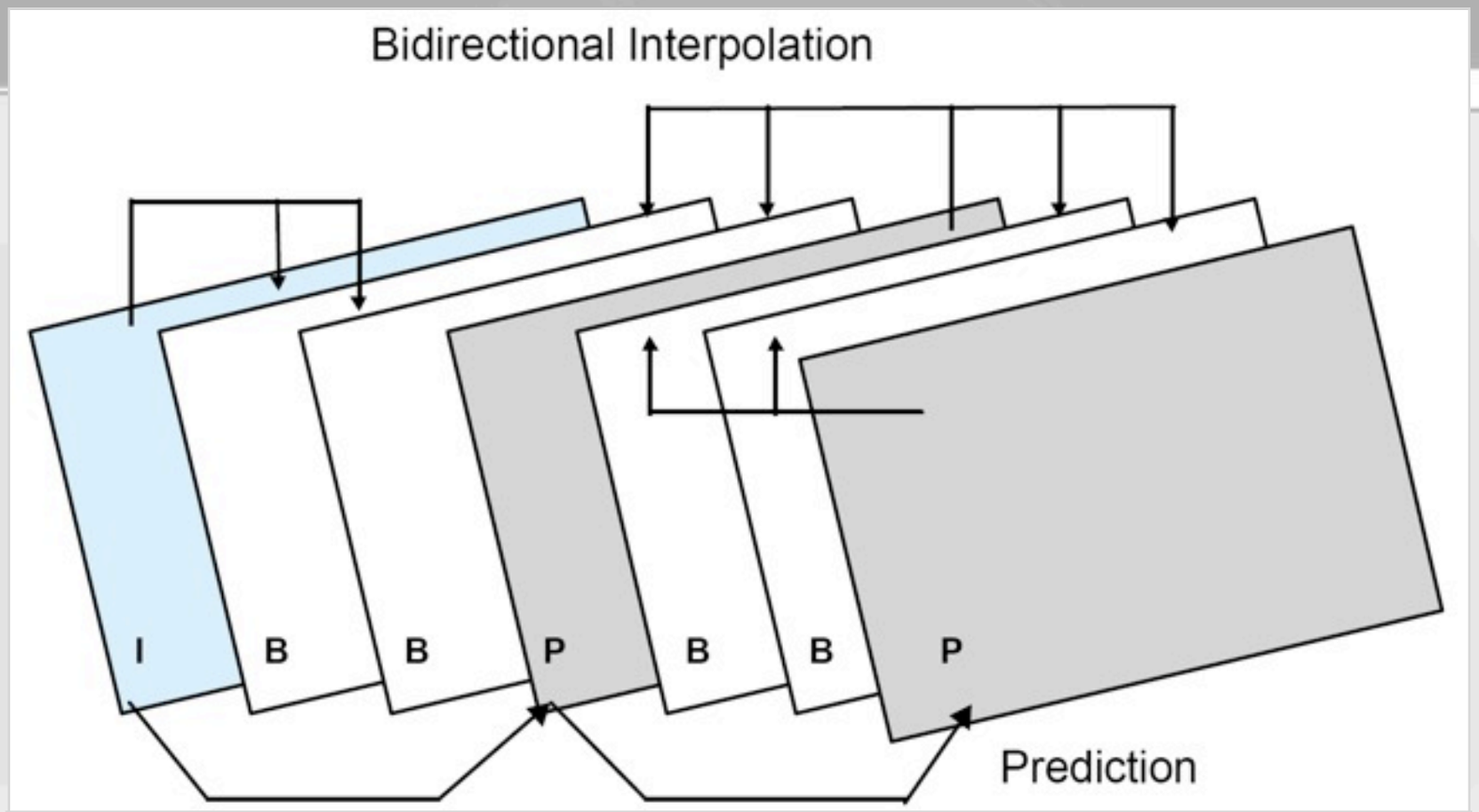
- 8 x 8 pixel block can be numerically manipulated by fast signal processor in real time
- Motion estimation is based on comparing the blocks between series of pictures

MPEG: only compress moving parts



MPEG: motion compression

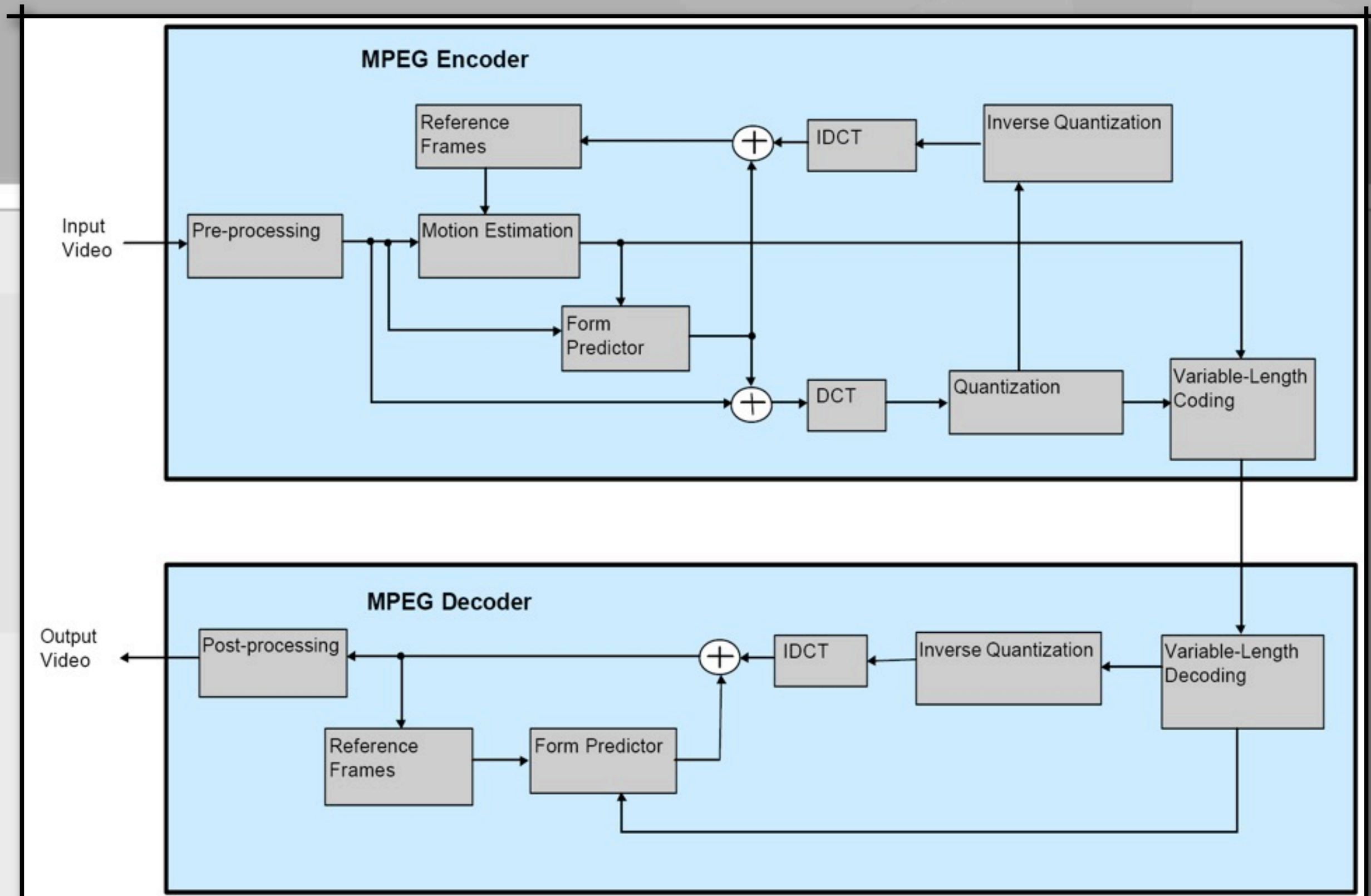
I = Intra-Frame
P = Predicted frame
B = Bi-directionally interpolated frame



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

MPEG: other issues

- Motion compensating
- Intra-frame transfer order

Color video coding

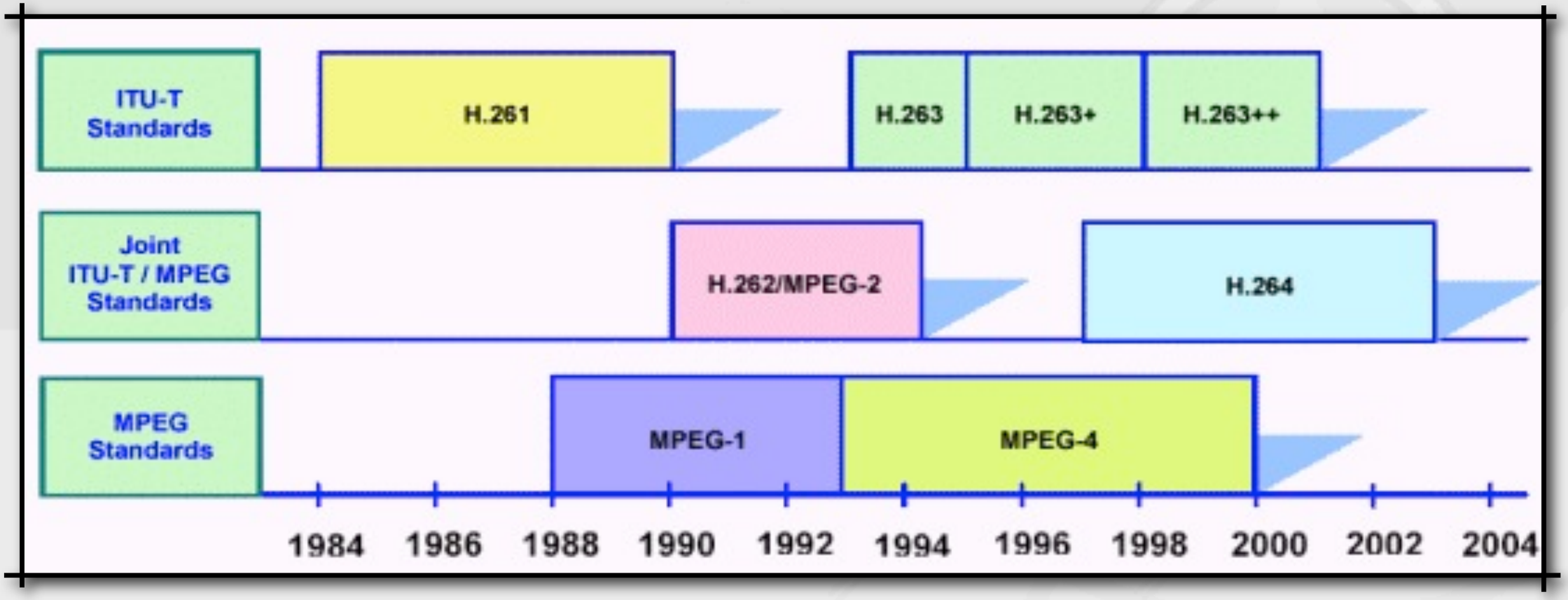
- 抽样和编码整个模拟（彩色）视频信号
  - 例如，复合编码
- 对亮度和色度分别编码
  - 例如，组件编码
  - 亮度比色度更重要，可根据应用场合采用4:2:2，4:2:0, 4:4:4等不同的编码比率

Overview of H.264

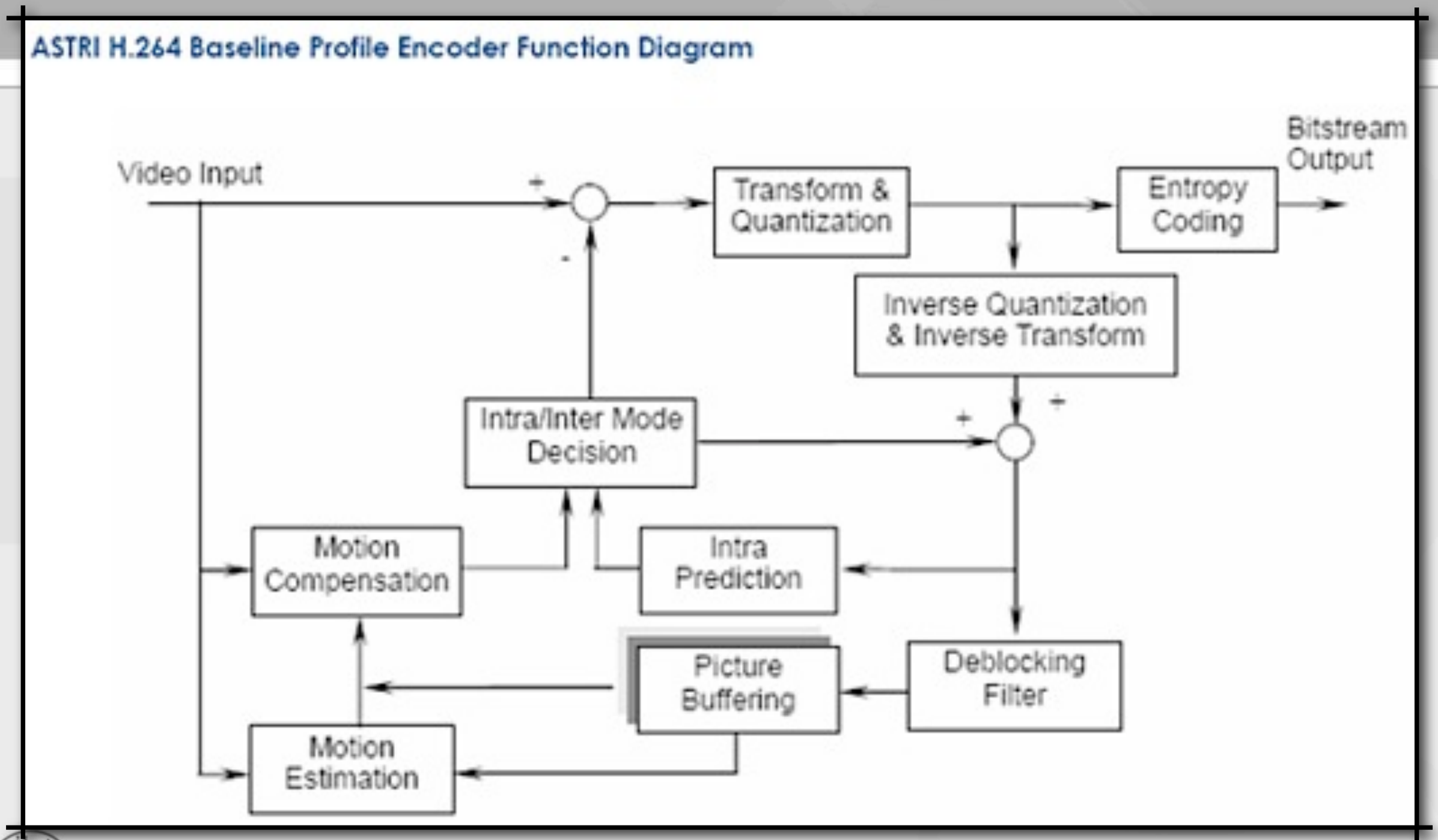
- JVT（Joint Video Team）
  - founded on December 2001, Pattaya Thailand.
  - video coding specialists from ITU-T and ISO, the two international standards organizations
  - goal: define a new video coding standards to achieve high compression rate, high image quality, good network adaptive coding frame.

- 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

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

2.4. HTML and XML

结构化文档概览

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 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

About CSS

HTML 5.0 !!!

- 学习参考： http://www.w3school.com.cn/html5/index.asp
- 实例解释： http://directguo.com/html5
- 一套Web富客户端开发的工业标准
  - 许多新特性：内建的视频、音频标记，元素拖放功能
  - 最新的 Safari、Chrome、Firefox 以及 Opera 支持某些 HTML5 特性。Internet Explorer 9 也将支持

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

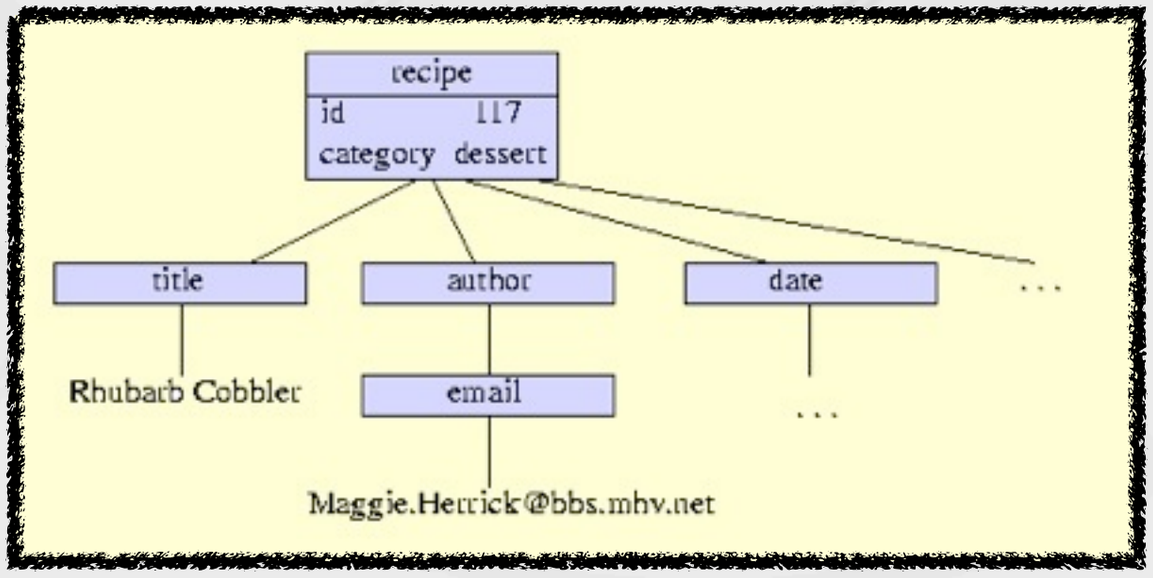
- An XML document is a (Unicode) text with markup tags and other meta-information.
- An XML document must be well-formed:
  - start and end tags must match
  - element tags must be properly nested
  - + some more subtle syntactical requirements

- 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,

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

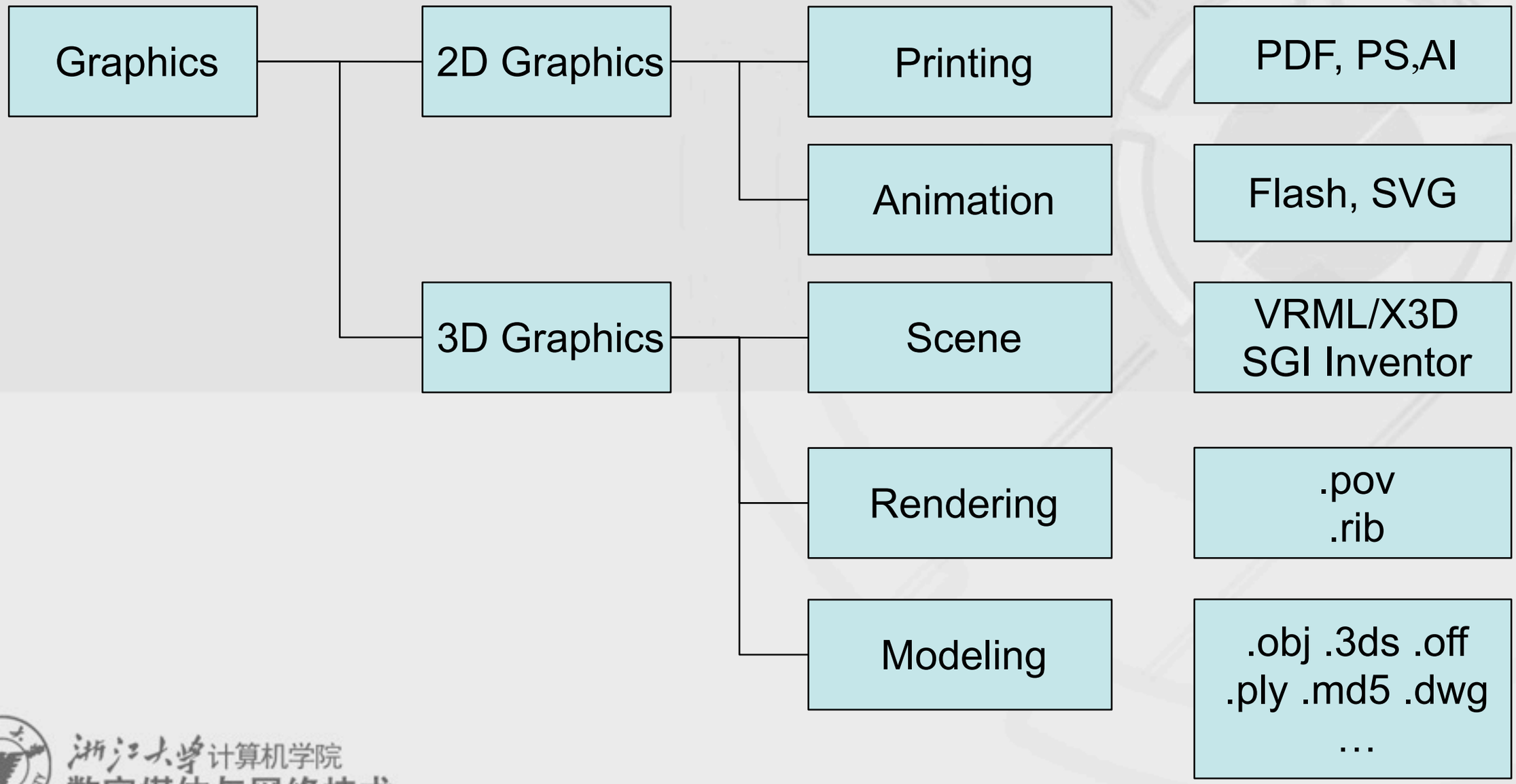
2.5. Graphics formats

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 SVG

- http://www.w3.org/Graphics/SVG/About.html
- 什么是SVG?
  - SVG 指可伸缩矢量图形 (Scalable Vector Graphics)
  - SVG 用来定义用于网络的基于矢量的图形
  - SVG 使用 XML 格式定义图形
  - SVG 图像在放大或改变尺寸的情况下其图形质量不会有所损失
  - SVG 是万维网联盟的标准
  - SVG 与诸如 DOM 和 XSL 之类的 W3C 标准是一个整体

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

X3D 教程

- http://x3d.esoe.ntu.edu.tw/
- Use FreeWRL as a player

COLLADA (SONY)

- Wildly used in PS3/PSP games
- XML based, and similar to X3D

- http://www.khronos.org/collada/
- http://www.opencollada.org/home.html

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.