



数字媒体技术基础 ——数字媒体压缩技术

王晗

wanghan@bjfu.edu.cn



专题提纲

- 数据压缩基础
- 图像数据压缩算法
- 动态图像压缩技术和标准
- H.26x标准

视频、图像编码标准

——国际标准化组织



国际电信联盟



美国电影电视工程师协会

Society for Motion Picture and Television Engineers



DTV 数字电视 HDTV 高清数字电视

ATSC (北美) DVB (欧洲) ISDB (日本)

视频、图像编码标准

- 数字音频编解码标准

- MP3, MP4

- 数字图像压缩标准

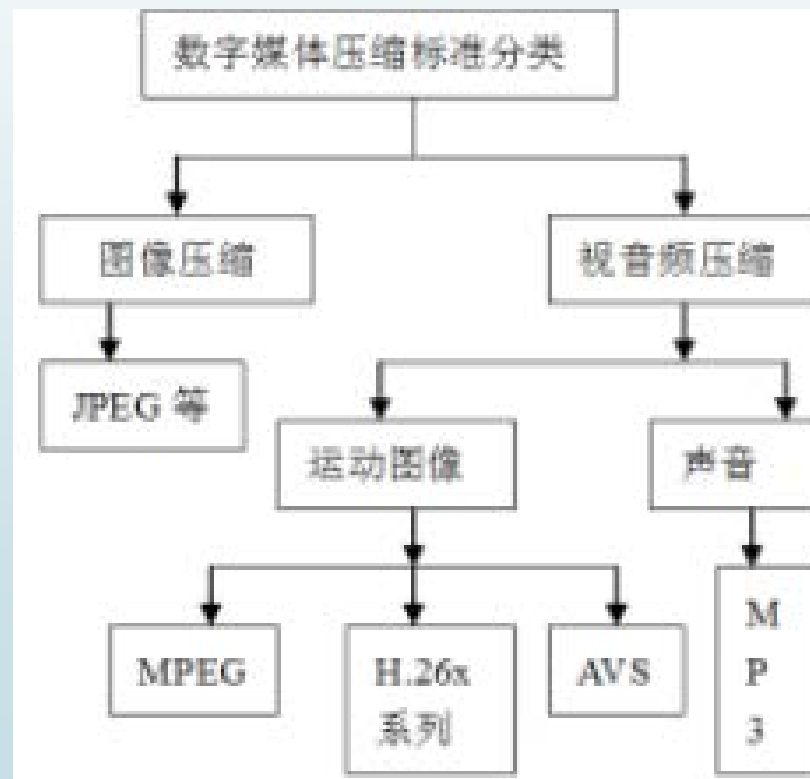
- JPEG

- 数字视频编解码标准

- MPEG系列

- H.26X系列

- AVS



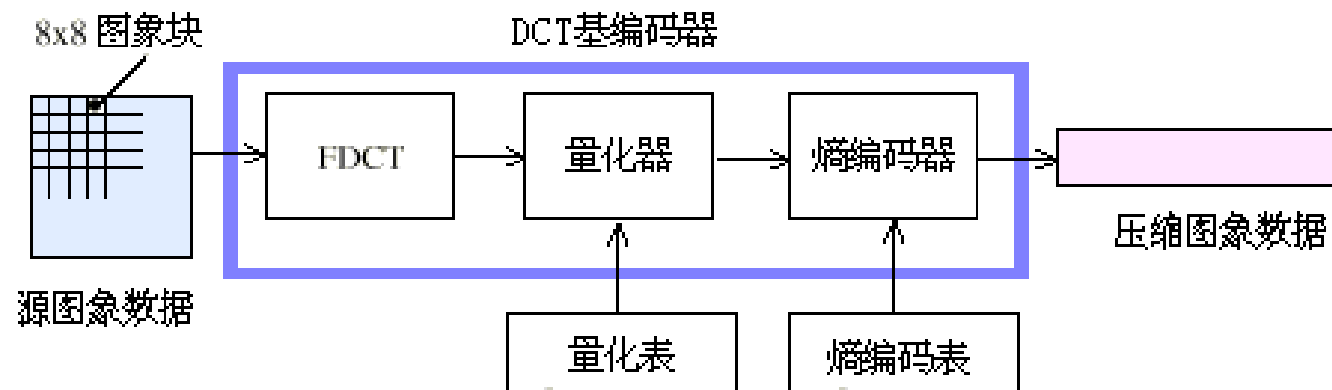
图像、视频编码标准

——国际标准化组织

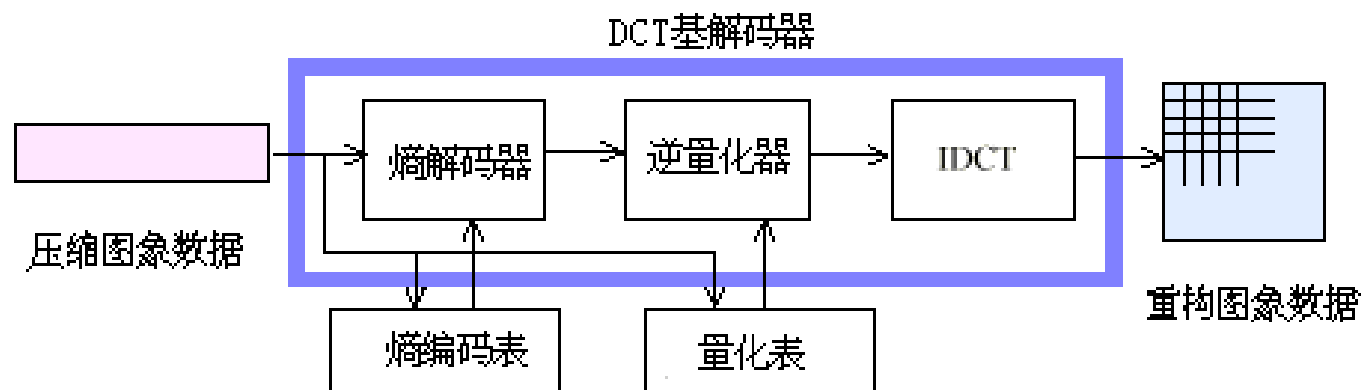
- 国际上主要有两大组织专门进行视频编码标准的制定工作
- ✓ ISO/IEC 下的JPEG(Joint Photographic Experts Group)
 - 一套国际静态图像压缩标准：ISO/IEC 10918 – 1号标准 “多灰度连续色调静态图像压缩编码” 俗称为JPEG，以其优异的性能，该标准一直到当前仍被因特网、数码相机等很多领域广泛应用。
- ✓ ISO/IEC 下的MPEG (Motion Picture Experts Group)
 - 1986年成立的运动图像专家组专门负责定制多媒体领域的相关标准，主要应用于存储、广播电视、因特网或无线网上的流媒体等
- ✓ ITU 下的VCEG (Video Coding Experts Group)
 - 主要制定面向实时视频通讯领域的视频编码标准，如视频电话、视频会议等应用

JPEG算法压缩编码

- JPEG包含两种基本的压缩算法：无损压缩算法（基于差分脉冲调制DPCM）和有损压缩算法（基于离散余弦变换DCT变换）。
- 右图为JPEG算法框图



(a) DCT基压缩编码步骤



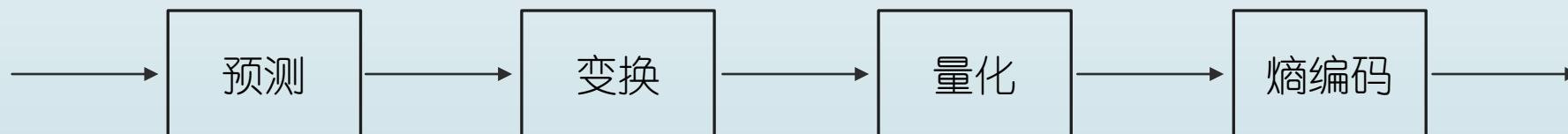
(b) DCT基解压缩步骤

JPEG算法压缩编码

- (1) 使用正向离散余弦变换 (Forward Discrete Cosine Transform, FDCT) 把信息从空间域变换成频率域的数据，并利用数据的频率特性进行处理；
- (2) 使用加权函数对DCT系数进行量化，这个加权函数对于人的视觉系统是最佳的；
- (3) 使用霍夫曼可变字长熵编码器对量化系数进行编码。

视频编码技术

- 视频压缩主要是通过去除视频中的空间冗余、时间冗余和编码冗余实现
- 视频编码器中包括很多编码算法，这些算法在编码器中被有效组合在一起，使整个编码器具有较高的压缩效率。
- 目前主流的视频编码器采用的技术主要有：预测、变换、量化、熵编码和环路滤波

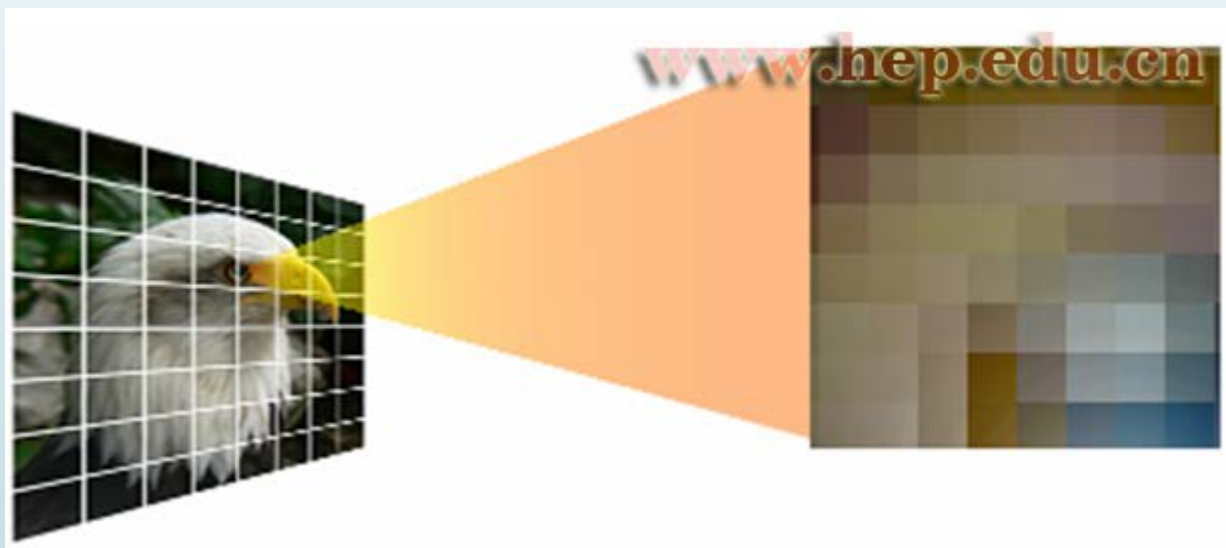


视频编码技术

- 视频压缩方法包括帧内压缩和帧间压缩
- 帧内压缩主要利用图像压缩技术，对每帧图像进行独立编码。其编码过程不依赖于前后帧图像数据。
- 帧间压缩主要利用数据的时域冗余，常用的方法有运动补偿变换编码。
- 主流的编码方法都是将图像划分成块进行编码，每一个块为一个编码单元，以编码单元为单位进行编码。
- 将每帧图像进行划分后，按照从上至下、从左至右的顺序对每个划分进行处理。



块可分为亮度块或色度块

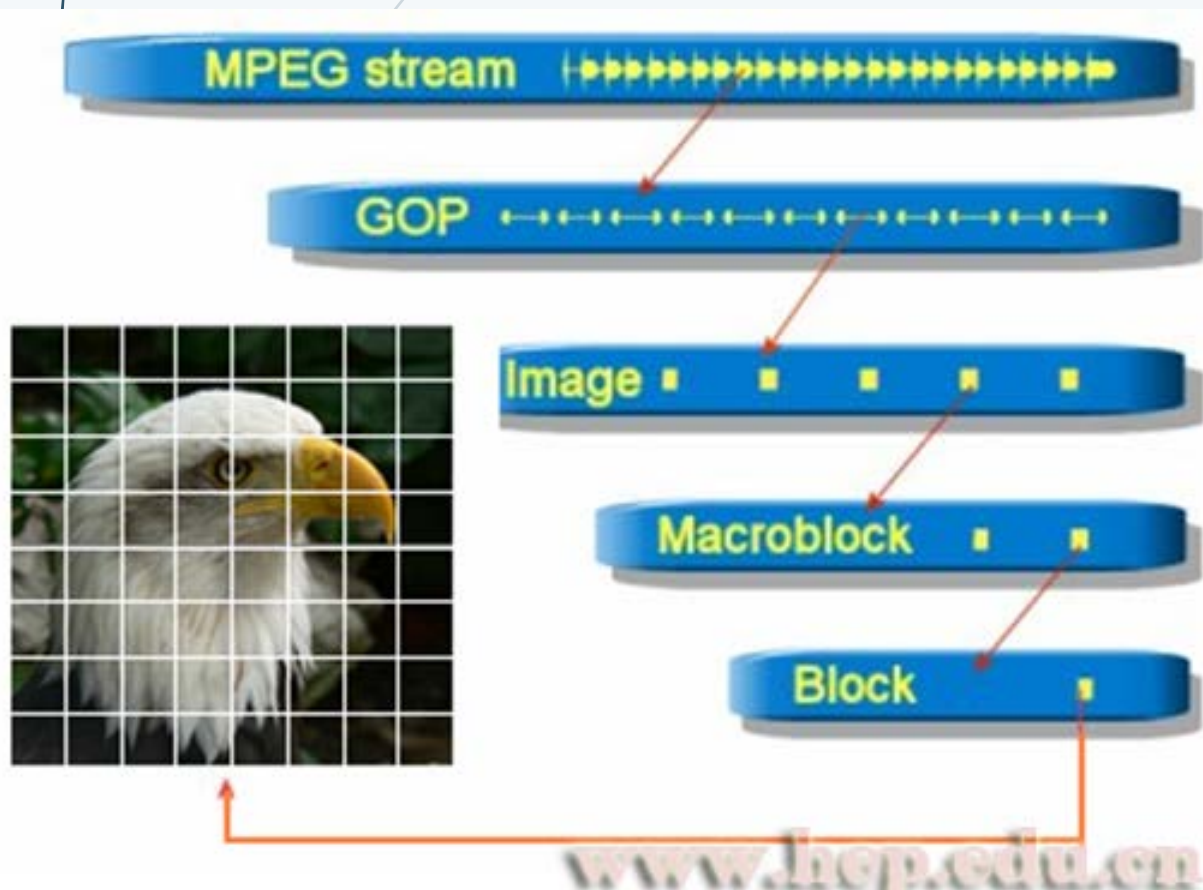


宏块是进行运动补偿的基本单位

视频编码标准

标准	名称	比特率
MPEG-1	面向数字存储的 运动图像 及其伴音编码	$\leq 1.5\text{Mbit/s}$
MPEG-2	运动图像 及其伴音信息的通用编码	$1.5\text{-}100\text{Mbit/s}$
MPEG-4	音视频对象的通用编码	$8\text{Kbit/s}\text{-}35\text{Mbit/s}$
H.261	$P \times 64\text{Kbit/s}$ 音视频业务编解码	$P \times 64\text{Kbit/s}(P:1\text{-}30)$
H.263	低比特率通信视频编码	$8\text{Kbit/s}\text{-}1.5\text{Mbit/s}$
H.264	先进视频编码	$8\text{Kbit/s}\text{-}100\text{Mbit/s}$

码流结构



► MPEG为更好地表示编码数据，规定了一个分层的结构，自上到下分别是：

- MPEG流 (MPEG stream)
- 图像组 (GOP, Group of Pictures)
- 图像 (Image)
- 宏块 (Macro block)
- 块 (Block)

视频编码标准

► MPEG-1

- MPEG-1标准 1992年公布，是MEPG工作组开发的第一个视频压缩标准
- 主要应用于数字媒体动态图像的存储和检索，如VCD
- 相关技术：采用了双向预测的B帧；自适应量化；逐行视频编码
- 编码码流分层结构：序列层-图像组层-图像层-条带层-宏块层-块层
- 首次定义了I、P、B三种图像类型的编码结构

视频编码标准

► MPEG-2

它是一个直接与数字电视广播有关的高质量图像和声音编码标准。MPEG2所能提供的传输率在3—10Mbit/s之间，在NTSC制式下的分辨率可达 720×486 ，可提供广播级的图像质量和CD级的音质。MPEG-2主要针对高清晰度电视(HDTV)所需要的视频及伴音信号，与MPEG-1兼容。

视频编码技术

——采样

- 采用YCbCr色彩空间可以有效地表达彩色视频图像，比RGB色彩空间更利于视频压缩
- YUV表示的亮度信号（Y）和色度信号（Cb、Cr）是相互独立的，可以对这些单色图分别进行编码。
- 根据HVS的特性，当彩色视频图像用YCbCr色彩空间表达时，可以用较高的分辨率表示Y分量，用较低的分辨率表示Cb和Cr分量
- 与RGB色彩空间相比，用YCbCr色彩空间表达视频图像可以减少所需的数据量，但对视频质量并没有明显的影响
- 视频编码中，常用YCbCr视频图像作为压缩对象，并对Y分量分配较多的编码比特数，对Cb和Cr分配较少的编码比特数。

视频编码技术

——预测

- 预测编码最早的系统模型是1952年贝尔实验室的Culter等实现的差值脉冲编码调制DPCM系统，其基本思想是不直接对信号进行编码，而是用前一信号对当前信号做出预测，对当前信号与预测值的差值进行编码传送。
- DPCM技术在视频编码中的应用分为帧内预测技术（消除空域冗余）及帧间预测技术（消除时域冗余）。

视频编码技术

——预测

- 帧内预测是利用图像在空间相邻像素之间具有很强相关性的特点
- 在现代视频编码中，采用了基于块的帧内预测技术
- 帧内预测利用图像在空间上相邻像素之间具有相关性的特点，由相邻像素预测当前块的像素值，可以有效地去除块间冗余

视频编码技术

——预测

- 帧间预测是消除运动图像时间冗余的技术
- Seyler在1962年发表的关于帧间预测编码的研究论文奠定了现代帧间预测编码的基础
- 视频相邻帧间存在很强的相关性，因此对视频序列编码只需编码相邻帧间的差异；相邻帧间的差异是由于物体的移动、摄像机镜头的摇动及场景切换等造成的
- 现代视频编码系统都采用了基于块运动补偿的帧间预测技术，用于消除时域冗余

视频编码技术

——预测编码

- 由于运动图像临近帧中的场景存在着一定的相关性，因此可为当前块搜索出在邻近参考帧中最相似的预测块，并根据预测块的位置，得出两者之间的空间位置的相对偏移量，即通常所指的**运动矢量**。
- 通过搜索得到的运动矢量的过程称为**运动估计**，根据运动矢量，从指定的参考帧中找到预测块的过程称为**运动补偿**。



前一帧画面（前一帧）

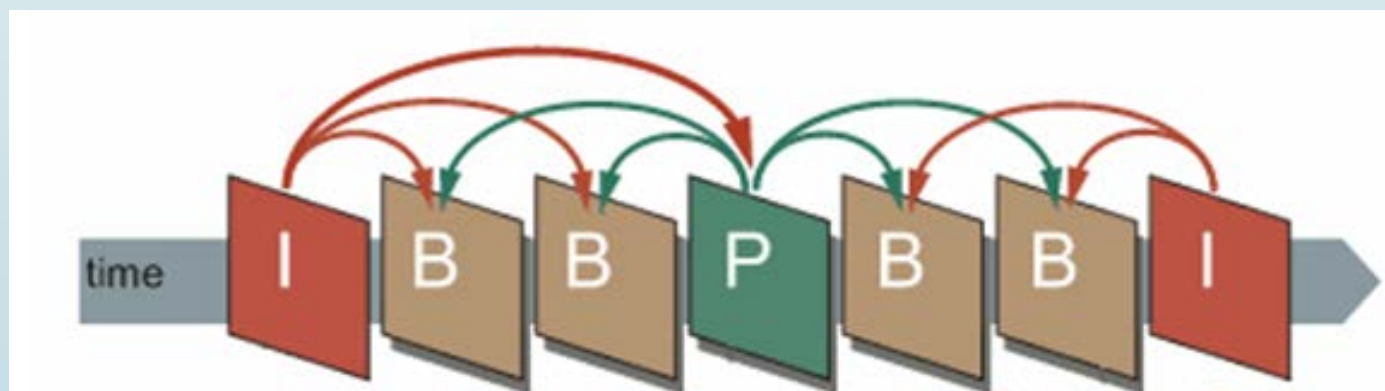


当前画面（当前帧）

视频编码技术

——预测编码

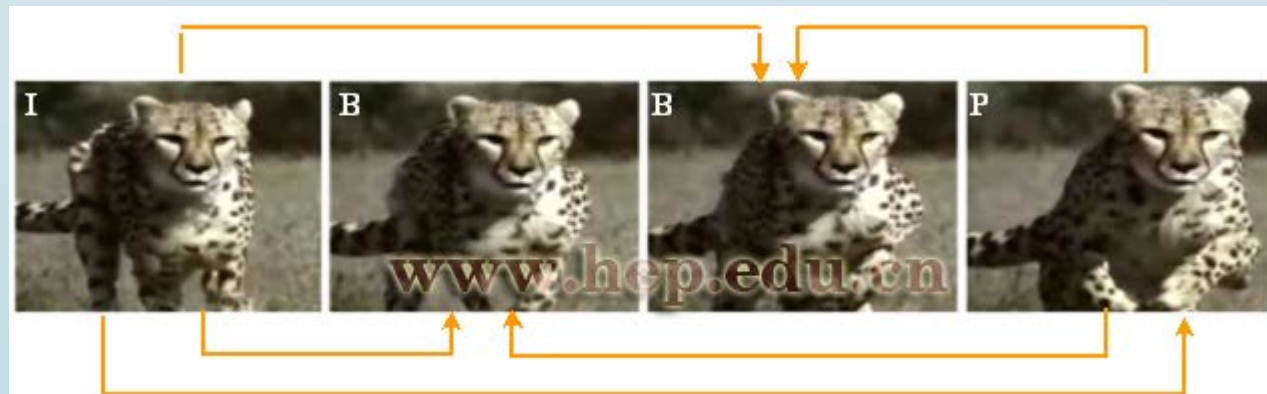
- 为了提高帧间预测的精度，基于块的运动补偿方案又从多个方面进行了完善。
- MPEG-1标准指定过程中发展出双向预测技术，即当前的预测值可以同时从前向参考帧和后向参考帧获得。
- 一个典型的MPEG帧序列



视频编码技术

——预测编码

- MPEG由三种类型的图像组成
- I图（帧内图Intra Picture）是对整幅图像采用JPEG编码的图像，它是一个独立的帧，其信息由自身画面决定，不需要参照其他画面而产生，它是P图和B图的参考图。
- P图（前向预测帧Predicted Picture），它参照前一幅I或P图像做运动补偿编码。
- B图像(双向预测 Bidirectional Prediction)，它参照前一幅和后一幅I或P图像做双向运动补偿编码。



视频编码技术

——预测编码

- 一个MPEG在解码播放时的实际帧序列



I帧

B帧

B帧

P帧

B帧

B帧

I帧

www.hep.edu.cn

视频编码技术

——变换编码

- 变换编码首先对图像进行正交变换以去除空间像素之间的相关性，也就是变换后的频域系数使图像信息更加紧凑的表示，这有利于编码压缩
- 正交变换使得原先分布在每个像素上的能量集中到频域的少数几个低频系数上（低频区域）
- 正交变换
 - K-L (Karhunen-Loeve)变换
 - DCT (Discrete Cosine Transform)变换
 - DWT (Discrete Wavelet Transform)变换

视频编码技术

——量化

- 量化是降低数据表示精度的过程，通过量化可以减少需要编码的数据量，达到压缩数据的目的。
- 量化可以分为矢量量化和标量量化两种
- 量化是一种有损压缩技术，量化后的视频图像不能进行无损恢复，因此导致源图像与重建图像之间的误差，称为失真。
- 通过在量化阶段调整量化步长，可控制视频编码码率和编码图像质量，根据不同应用的需要，在两者之间进行选择 and 平衡

视频编码技术

——熵编码

- 熵编码技术是视频编码系统中的基础性关键技术之一
- 目标是去除信源符号在信息表达上的表示冗余，也成为信息熵冗余或者编码冗余
- 负责利用信息熵原理进行数据的最终压缩
- 编码一个视频序列究竟需要多大的数据量？
- 信息量是什么？熵是什么？

基于对象编码技术的MPEG-4

► 传统MPEG编码技术的局限

- 在传统的视频压缩中，以一连串的影像(image)为压缩对象，也就是对于一整张画面(帧frame)做数据压缩。如此会将整张画面不重要的部分（如单调的背景）也一起压缩进去而占有一定程度的数据量。另外，将图像固定地分成相同大小的块，在高压缩比的情况下会出现严重的块效应，即马赛克效应。
- 未能考虑信息获取者的主观意义与人类视觉系统（HVS, Human Visual System）主观特性，未能考虑事件本身的特性如具体含义、重要性以及后果等等，不能对图像内容进行访问、编辑和回放等工作；
- 缺乏一套有系统与有效率的标准。

视频编码标准

► MPEG-4

MPEG-4旨在为视音频数据的通信、存取与管理提供一个灵活的框架与一套开放的编码工具。

不是一个单纯的音视频编解码标准，它将内容与交互性作为核心，从而为多媒体数据压缩提供了一个更为广阔的平台

提供基于对象的编码功能，注重于用户与内容的交互性，实现各种新工具以支持编码对象及视频之间的融合

基于对象编码技术的MPEG-4

► 何谓对象编码

- 基于对象编码，是指依赖于对视频场景中任意形状对象检测的编码机制。MPEG-4是目前基于对象编码机制，是对运动图像和伴音编码的算法。



(a)



(b)



(c)



(d)



(e)



(f)

基于对象编码技术的MPEG-4

■ MPEG-4核心编码技术-AV对象（Audio/Visual Objects）

- AV对象是MPEG-4为支持基于内容编码而提出的重要概念。对象是指在一个场景中能够访问和操纵的实体，对象的划分可根据其独特的纹理、运动、形状、模型和高层语义为依据。

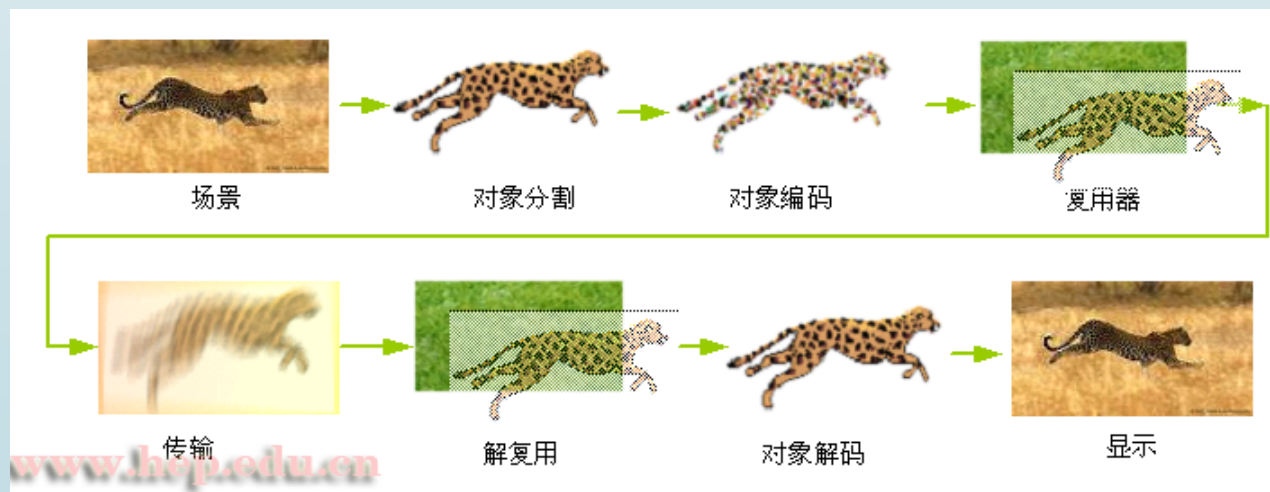
■ MPEG-4对AV对象的操作主要有：

- 1. 采用AV对象来表示听觉、视觉或者视听组合内容。
- 2. 允许组合已有的AV对象来生成复合的AV对象，并由此生成AV场景。
- 3. 允许对AV对象的数据灵活地多路合成与同步，以便选择合适的网络来传输这些AV对象数据。
- 4. 允许接收端的用户在AV场景中对AV对象进行交互操作。
- 5. MPEG-4支持AV对象知识产权与保护。

基于对象编码技术的MPEG-4

■ 基于场景的对象分割与编码技术

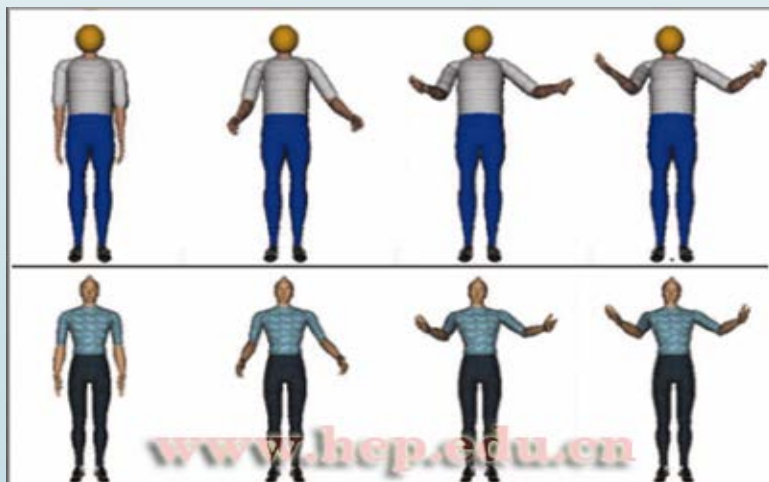
- MPEG-4实现基于内容交互的首要任务就是把视频/图像分割成不同对象或者把运动对象从背景中分离出来，然后针对不同对象采用相应编码方法，以实现高效压缩。
- MPEG-4的编码理念是：在编码时将一幅景物分成若干在时间和空间上相互联系的视频音频对象，分别编码后，再经过复用传输到接收端，然后再对不同的对象分别解码，从而组合成所需要的视频和音频, 其过程如图所示。



基于对象编码技术的MPEG-4

► 视频对象平面-VOP (Video Object Plane)

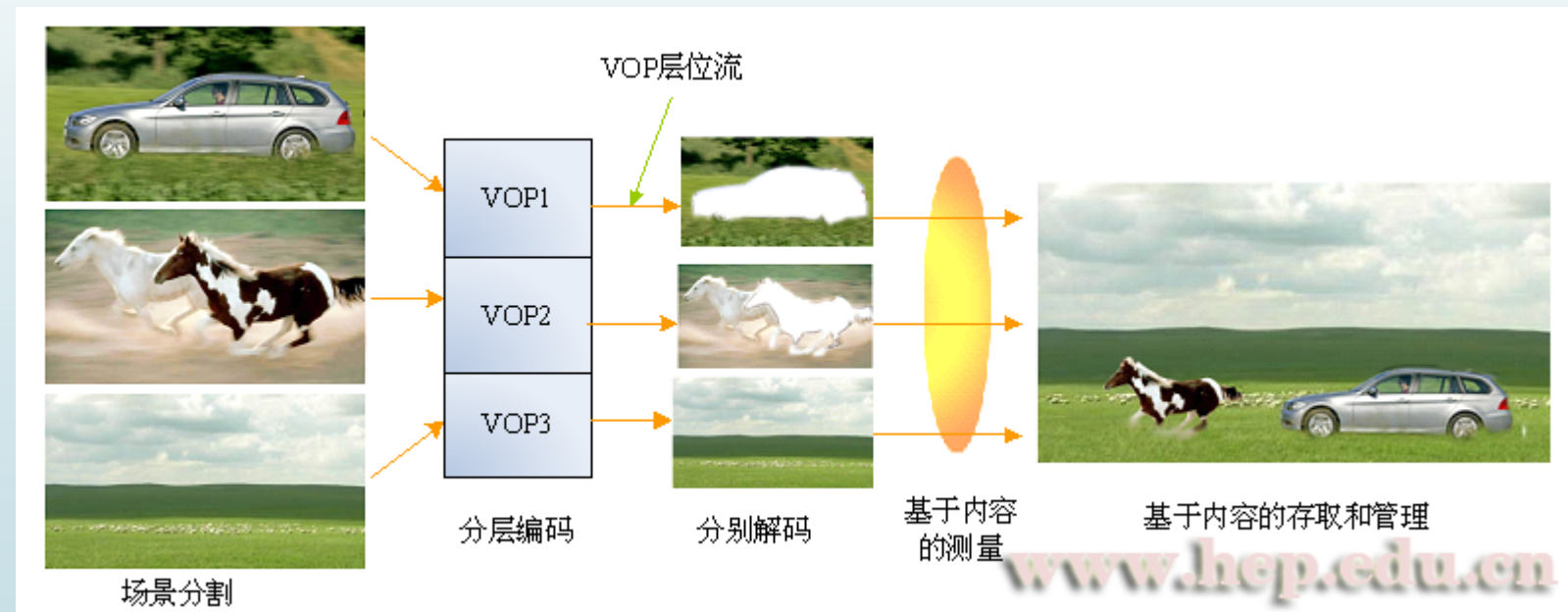
- 我们根据人眼感兴趣的一些特性如形状、运动、纹理等，将图像序列中每一帧中的场景，看成是由不同视频对象平面VOP (Video Object Plane)所组成，而同一对象连续的VOP称为视频对象VO。
- VOP是视频场景的语法对象，除轮廓信息外，主要由亮度和色度分量(Y, U, V) 组成。VOP可以是任意形状，可由半自动分割来检测。



上图为实现基于内容交互，人物由简单到复杂的轮廓物征提取

基于对象编码技术的MPEG-4

- ▶ 视频对象平面 (VOP, Video Object Plane) 是视频对象 (VO) 在某一时刻的采样, VOP是MPEG-4视频编码的核心概念。



VOP视频编码技术

H.26x标准

- 鉴于读视频图像传输的需求以及传输带宽的不同，1990年和1995年制定了适用于综合业务数字网（ISDN）和公共交换电话网（PSTN）的视频编码标准，即H.261协议和H.263协议。
- H.264标准也称为MPEG-4 Part10——高级视频编码。
- H.264视频编码标准的主要目标和特点为：提高压缩编码效率和增强网络适应能力
- H.264关键技术：
 - 分层设计
 - 高精度、多模式运动估计
 - 4*4块的整数变换
 - 统一的VLC
 - 帧内预测
 - 切换帧（SP和SI）
 - 面向IP和无线环境

AVS标准

- AVS (Audio Video Standards) 是中国自主制定的音视频编码技术标准，其核心是把数字视频和音频数据压缩为原来的几十分之一甚至百分之一以下。
- AVS标准包括系统、视频、音频、数字版权保护等四个主要技术标准和一致性测试等支撑标准。
- AVS-视频当中具有特征性的核心技术包括：8x8整数变换、量化、帧内预测、1/4精度像素插值、特殊的帧间预测运动补偿、二维熵编码、去块效应环内滤波等。

随堂实验

- 参考实验指导材料《python视频处理入门》，完成材料中给定的五个练习
- 提交材料: 1.实验报告
- 提交截止时间：4月23日；
- 提交文件夹：ftp://211.71.149.149/Wang_Han/homework/数字媒体技术基础/数媒17/数字视频处理/0409
- 命名要求：学号_姓名.doc