## 数据压缩原理与应用

张远

yzhang@cuc.edu.cn

### 课程基本情况

- 专业课,48学时/16学时授课+32学时实验
- ■先修课程
  - ▶ 高等数学、线性代数、概率论与数理统计、信息论与编码原理,信号与系统,数字信号处理
  - C/C++/python...
- ■上课时间
  - ▶ 单周周一上午34节
  - > 双周周一上午1-4节

### 课程基本情况

- ■教材
  - 》数据压缩(第三/四版) 吴乐南,电子工业出版社
  - >自编辅助讲义(随课程进度发放)

■参考教材

Introduction to Data Compression数据压缩导论(第4版) 人民邮电出版社

- Khalid Sayood

### 教学要求 (安排)

- ■课堂听讲
  - >课堂考勤;课堂秩序
- ■实验
  - > 提高分析和动手能力的重要环节
  - > 认真撰写实验报告,提高归纳表达能力
- ■答疑、其他交流方式
  - > 电子邮件、实验报告,课程小结等
  - > 随时,只要在线(线上教学期间)

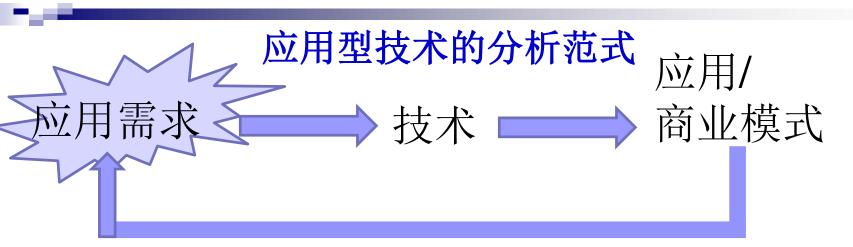
### 考核方法

- 平时成绩 + 期末考试
  - 产平时成绩: 40%(考勤、作业、实验报告)
  - ▶期末考试: 60%(基础知识和实验内容书面考试)

# 第一章绪论

### 第一章: 绪论

- ■数据压缩问题的分析
- ■数据压缩技术的基本模型
- ■数据压缩应用中的问题
- ■数据压缩系统的性能评价
- ■本课程的主要内容
- ■本课程的教学目标



减少容纳给定消息集合或数据采样集合的信号空间

- > 空域: 物理存储空间
- >时域:时间区间
- >频域: 传输带宽
  - □不同调制方式——频带压缩

编码的对象? 无损 or 有损

压缩?

Introduction



#### 编码(压缩)的对象:

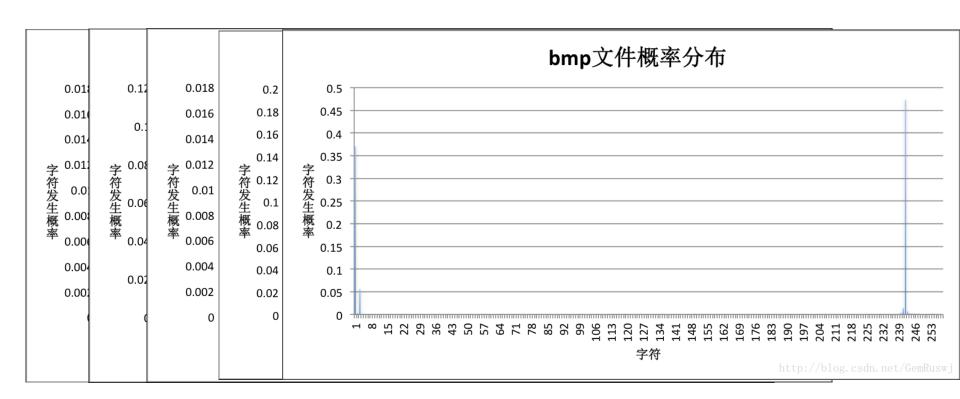
- > 数据
- ▶ 图像/计算机图形
- > 视频
- > 动画
- ▶语音
- > 音频...

数据的特点是什么?

是否存在通用性技术?

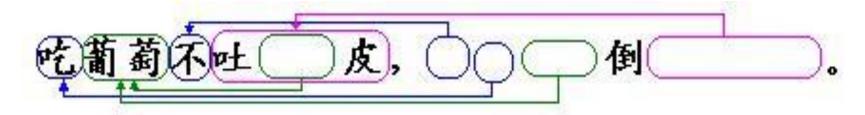
数据的特点:是什么?

> 包罗万象,各种可能(文档/文件夹、采集数据等等)



信息论:统计冗余(剩余度)——统计匹配编码(两遍编码)

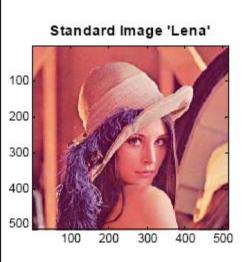
数据的特点: 如果事先不可能知道概率分布怎么办?

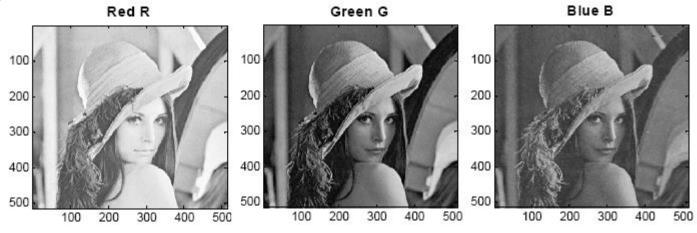


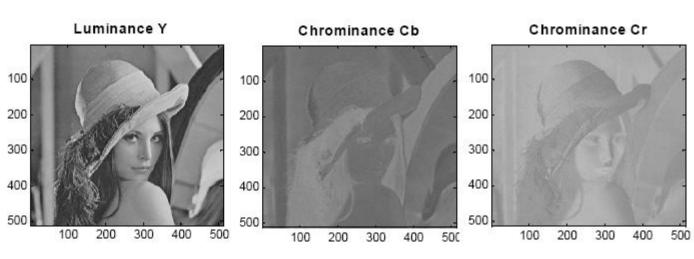
#### 通用编码: 词典编码

- ▶ 词典编码的根据是数据本身包含有重复代码这个特性。 例如文本文件和光栅图像就具有这种特性。
- ▶ 基本思想:构造一个词典,将数据中反复出现的字符串 ,用(词典中的)索引号表示,解码时通过查词典,转 换为原字符串。
- > 实现时的问题?
- Winzip, rar....

#### 图像的特点:



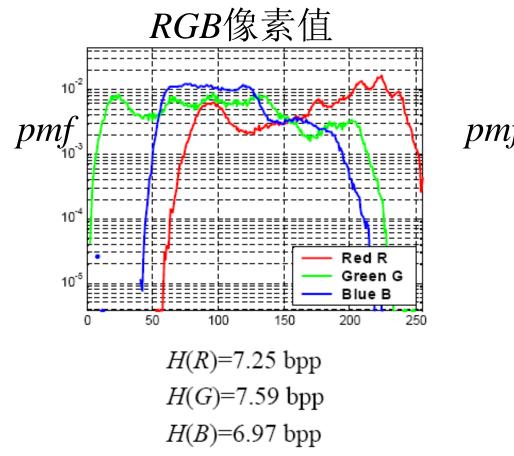


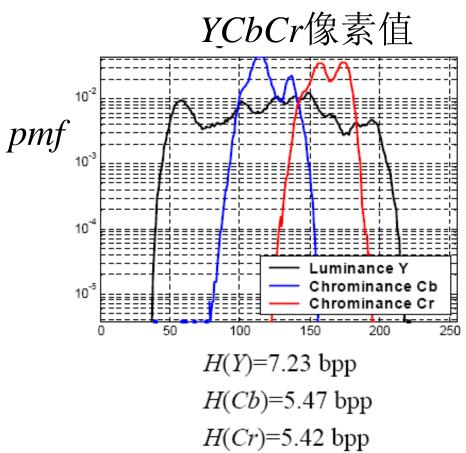


> 原始图像数据空间的分析

Introduction

#### 图像的特点:





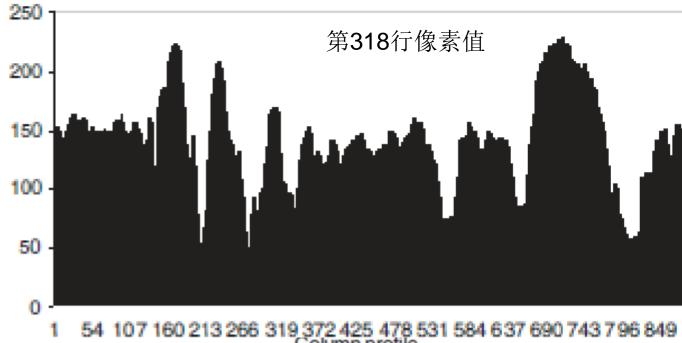
> 原始图像数据空间的分析

Introduction

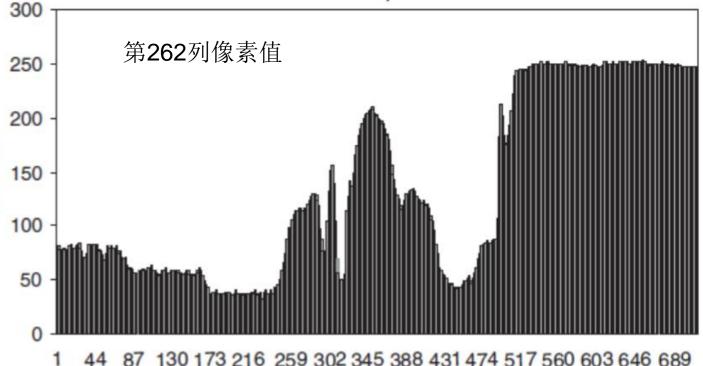
### 图像的特点

Gray level



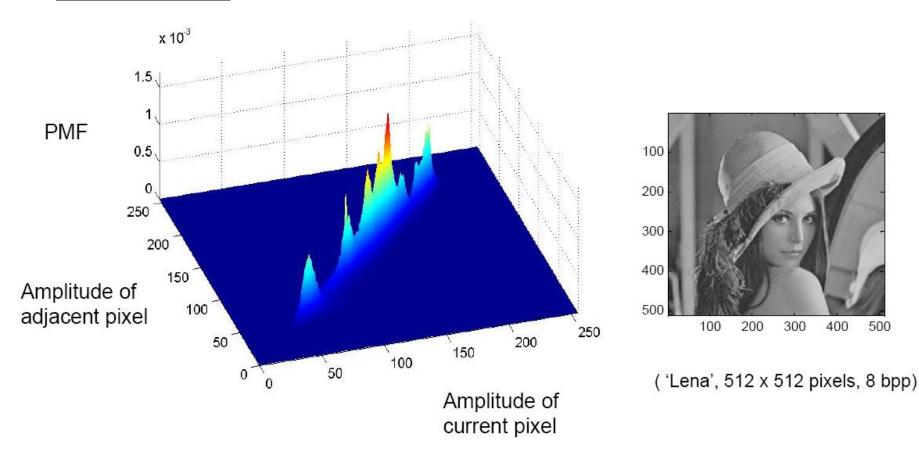


54 107 160 213 266 319 372 425 478 531 584 637 690 743 796 849 Column profile



1-14

#### 图像的特点:

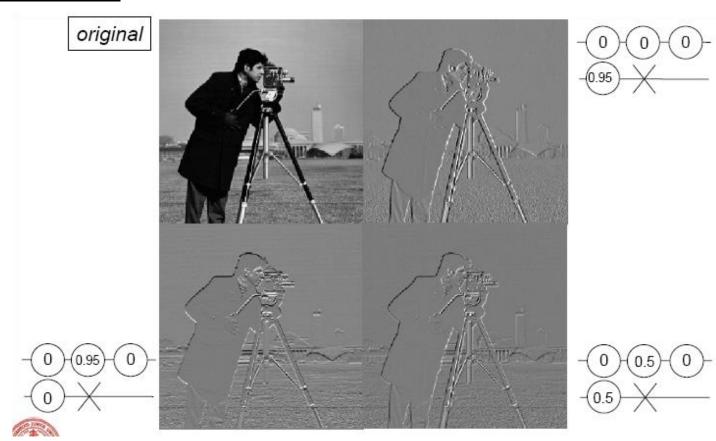


> 大量的空间冗余。如何利用?

Introduction 1-15

500

#### 图像的特点:



> 进一步分析预测误差图像和原图像的数据特点?

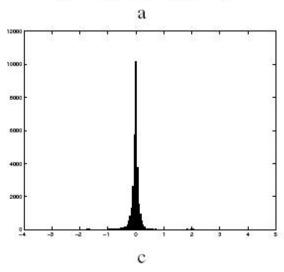
1-16 Introduction

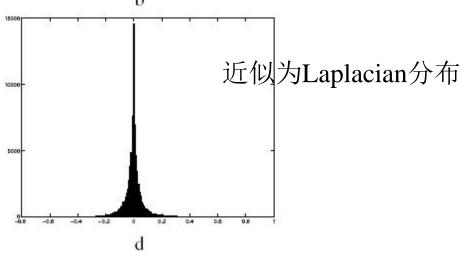
#### 图像的特点:

原始图像

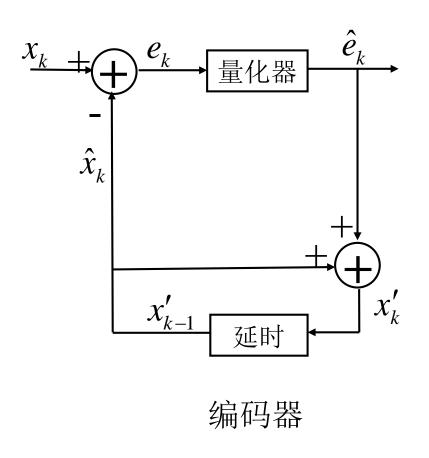


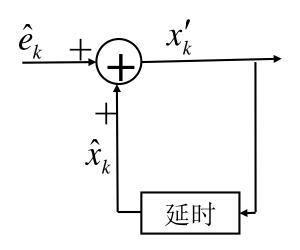






预测+熵编码(huffman编码)——算法组合是个利器!





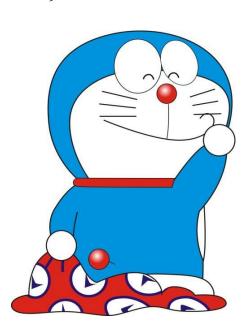
解码器

算法组合带来了哪些问题? 有无优化方法?

■ 矢量图形是用一系列计算机指令来表示一幅画,如所有直线,曲线,圆,圆弧,矩形等图形元素。这种方法实际上是用数学方法来描述一幅图。

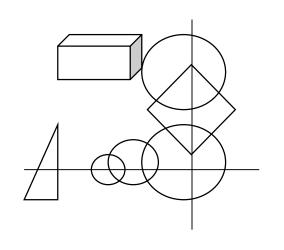
■ 在屏幕显示矢量图形时,首先将描述图形的指令转化成屏幕上的形状和颜色。编辑矢量图的软件通常称为绘图软件,

如AutoCAD.

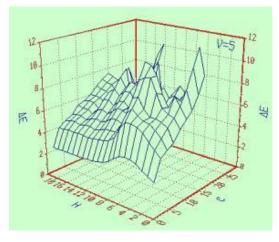




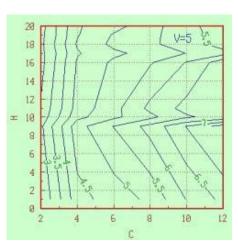
#### 计算机图形 (矢量图) 的特点:



(a) 矢量图形举例一



(a) 矢量图形举例二

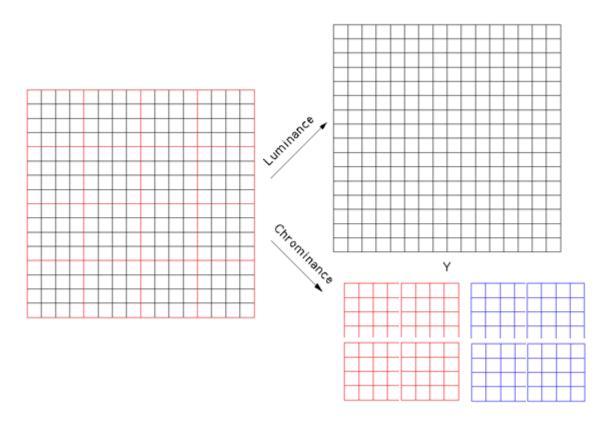


(b) 矢量图形举例三

- > 文件小;
- > 图像元素对象可编辑;
- > 图像放大或缩小不影响图像的分辨率;
- > 图像的分辨率不依赖于输出设备。

Introduction

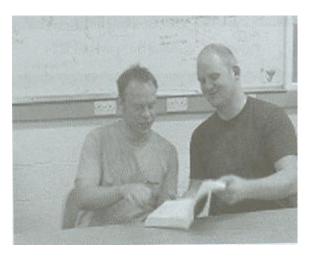
### 视频的特点:



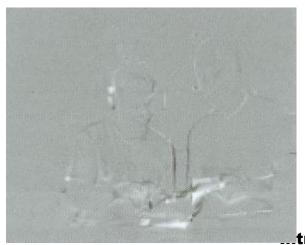
通常针对传输采用的是4:2:0采样格式。

#### 视频的特点:空间/时间的冗余









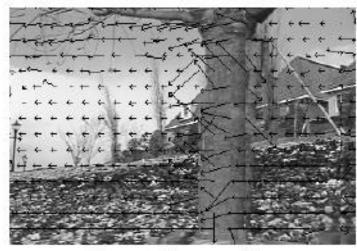
### 1.1 数据压缩问题的分析-运动估计预测

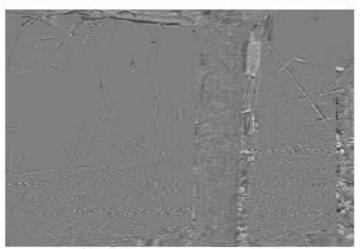
#### Current frame







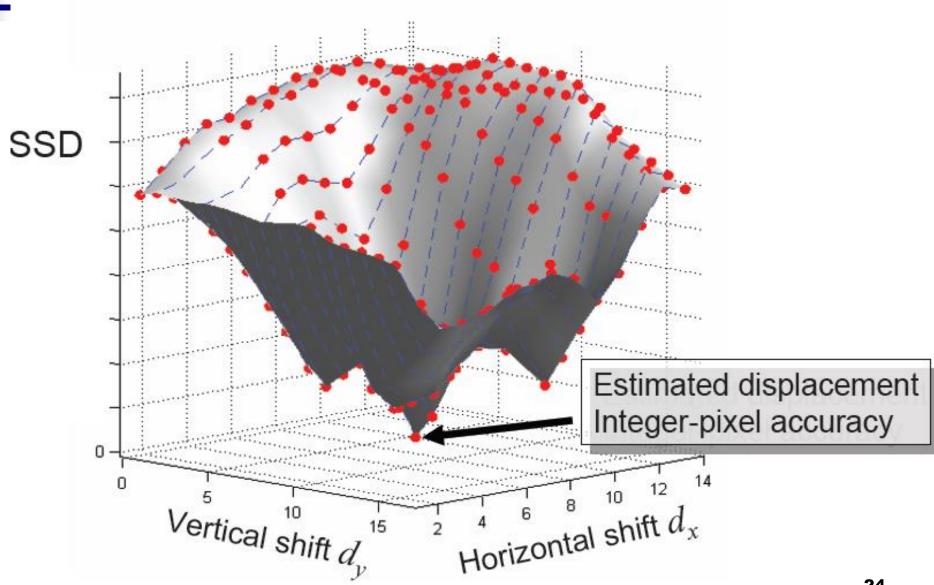




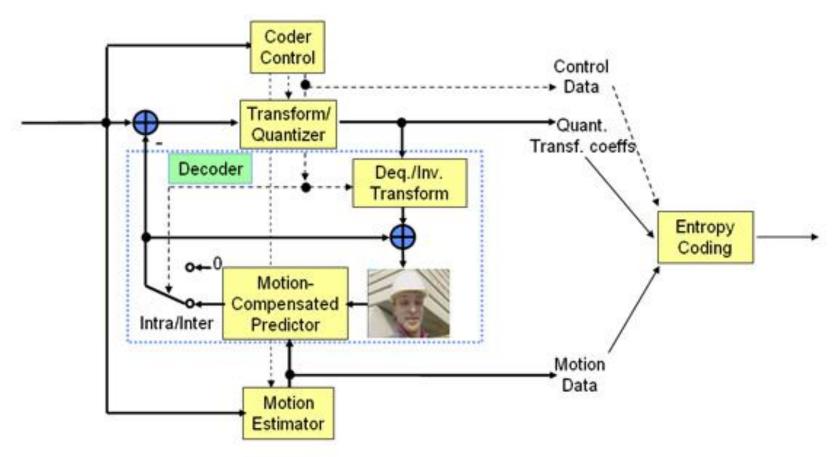
Current frame with displacement vectors

Motion-compensated Prediction error

### 1.1 数据压缩问题的分析-运动估计预测



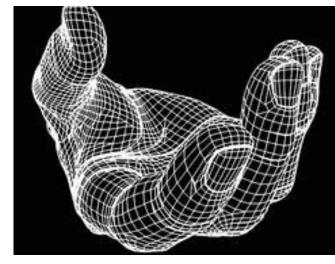
视频压缩的思路: 算法组合—预测+变换+量化+熵编码



- > 算法各个模块对总体性能的贡献?
- > 如何优化设计一个算法组合体?

#### 动画/游戏图形的特点:





- > 三角网格数据的压缩
  - 几何信息:每个顶点在三维空间中的具体位置、颜色、法向、纹理坐标;
  - ▶ 拓扑信息:各分量的数目以及相互之间的连接关系。

Introduction

#### 动画/游戏图形的特点:

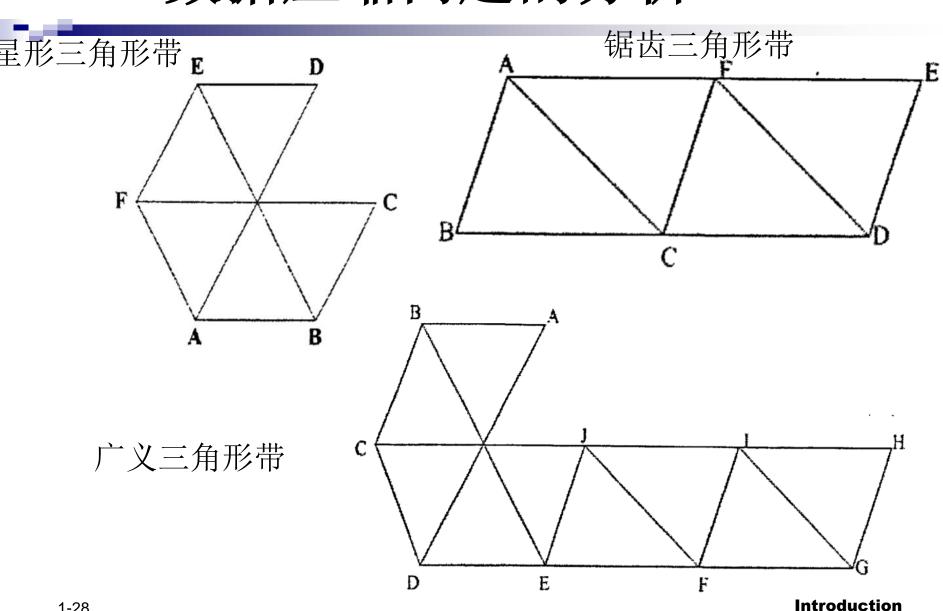
- > 三角网格数据的数据结构表示(以VRML ASCII格式 为例)
  - > 用一个坐标数组存储三角网格中所有顶点坐标
  - 用一个面片数组列出每个三角形面和它的三个顶点对应关系。
  - ➤ 一个由n个顶点,m个面组成的三角网格可以表示 为顶点坐标的集合V和顶点连接信息F的二元组

M={V,F}, 其中
$$V = \{V_1, V_2, \dots, V_n\}$$
;  $F = \{F_1, F_2, \dots, F_m\}$ 

每个顶点坐标  $V_i = [x_i, y_i, z_i]$ 

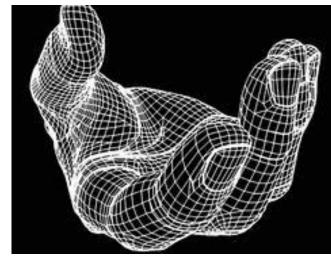
$$F_i = [V_a, V_b, V_c]$$
,其中 $0 \le V_a, V_b, V_c \le n$ 

ntroduction



#### 动画/游戏图形的特点:

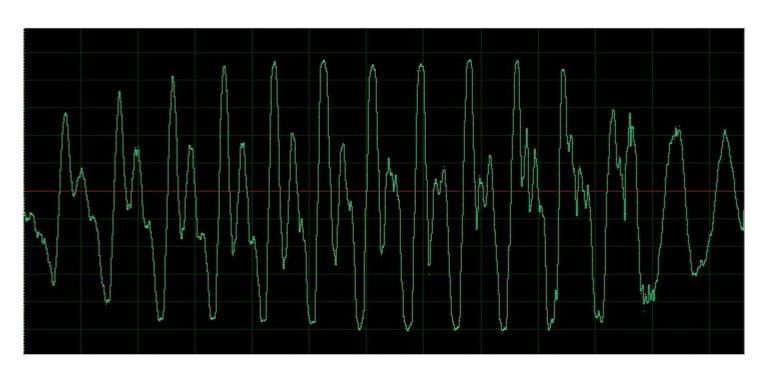




- > 三角网格数据: 拓扑信息和几何信息;
- > 纹理数据的压缩—参数化生成复杂的贴图图案;
- > 场景合成的数据结构表示—对3D空间进行划分, 对其中各种不同类型的节点进行统计分析

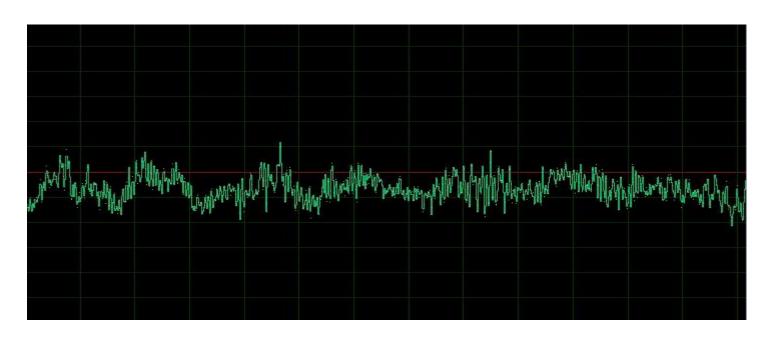
Introduction

### 语音数据的特点:



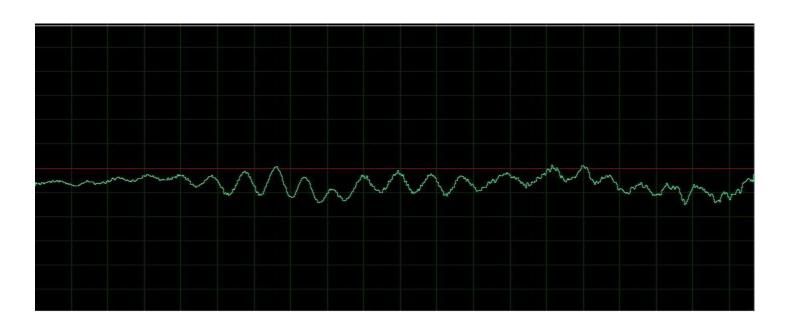
U的波形(浊音)

#### 语音数据的特点:

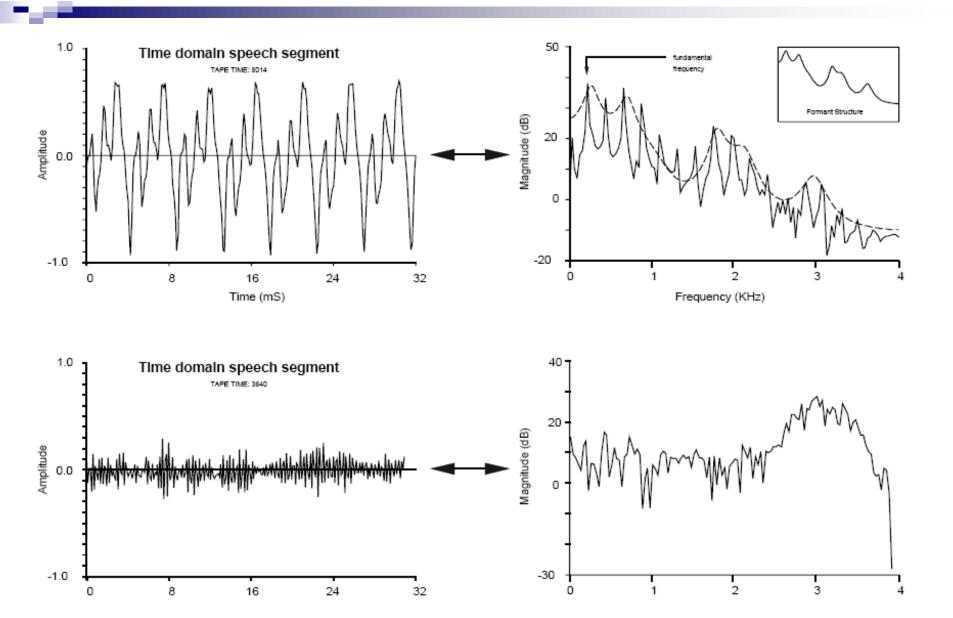


T的波形 (清音)

#### 语音数据的特点:



ao的波形



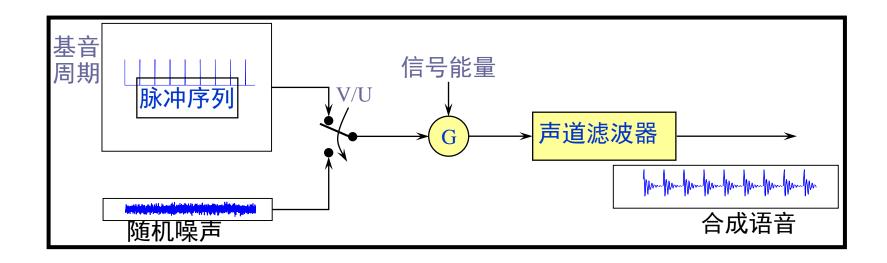
#### 语音数据压缩的思路:

- ➤ 语音数据的物理产生过程对于所有人来说都是一致的,即:肺部的气流冲击喉咙的声带,通过舌头和口型变化发出不同的声音。
- ➤ 如果能想办法从语音数据产生物理过程的建模入手 ,不直接传送语音信号本身,只传送其模型信息, 那么就可能实现很大程度的数据压缩。在接收端只 需要利用模型信息重建语音即可。



Introduction

#### 语音数据压缩的思路:

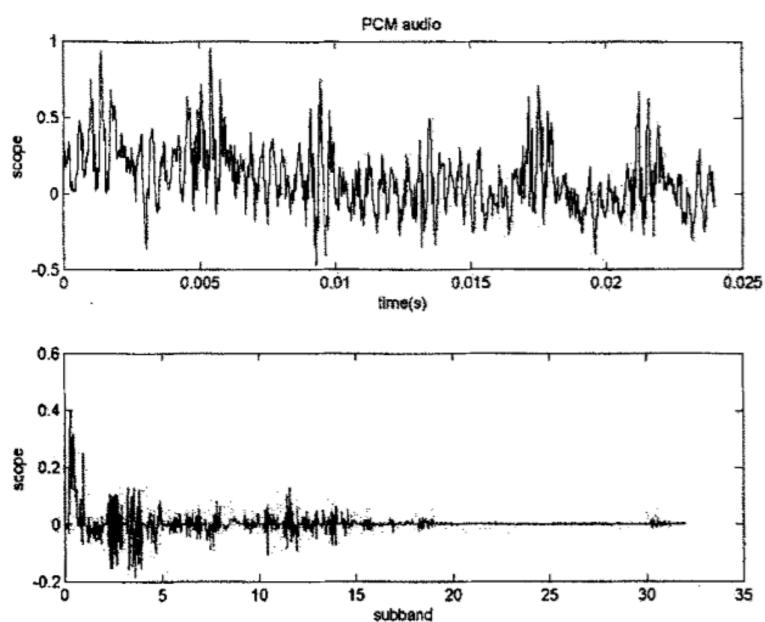


参数: 浊音/清音标记

基音周期

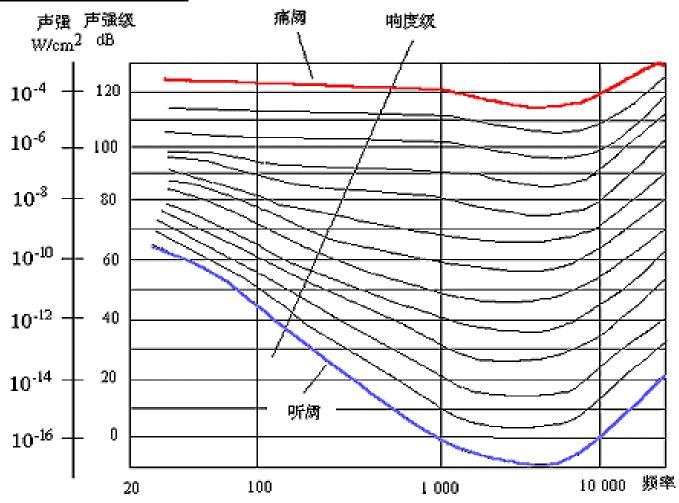
信号能量

音频数



### 1.1 数据压缩问题的分析

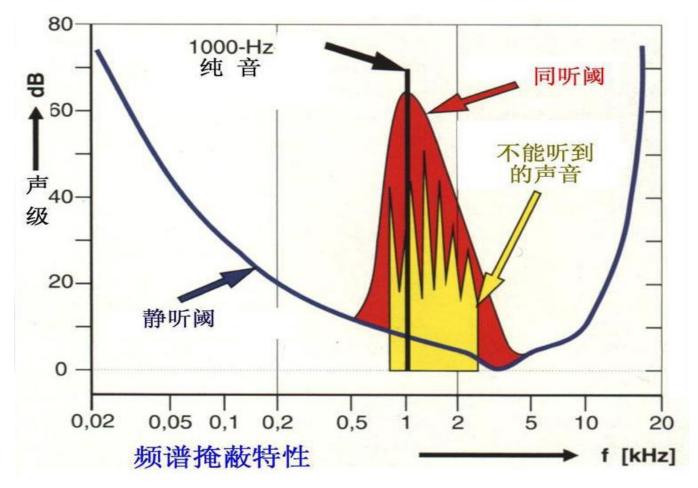
#### 人耳的感知特性:



人耳"听阈-频率"曲线

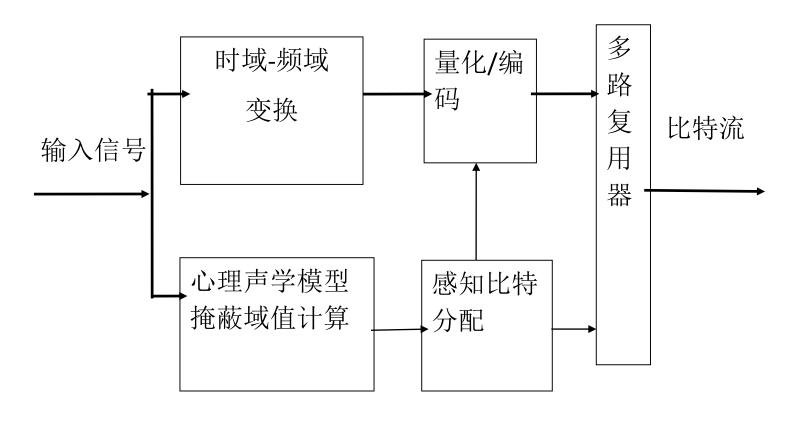
#### 1.1 数据压缩问题的分析

#### 人耳的感知特性:



#### 1.1 数据压缩问题的分析

#### 音频数据压缩的思路:

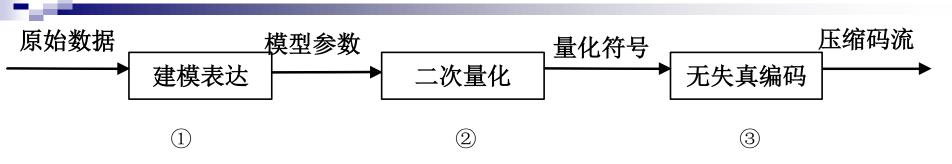


#### 感知音频编码基本框架

#### 1.2 数据压缩技术的基本模型

- 从信号的统计结构出发进行考虑(信息论为理论 分析的基础)
  - > 概率不均匀性
  - >相关性(记忆性):一维/二维/三维
- 从信源数据的物理产生过程出发进行考虑(随机信号处理为理论分析的基础)
  - > 如利用人类的发声系统,设计语音压缩算法
    - ■可用在军事、移动通信和玩具中的语音合成中
- 从人类的感知特性出发进行考虑 (信号处理理论)
  - > 听觉冗余——感知音频编码
  - ▶视觉冗余——图像/视频编码

#### 1.2 数据压缩技术的基本模型



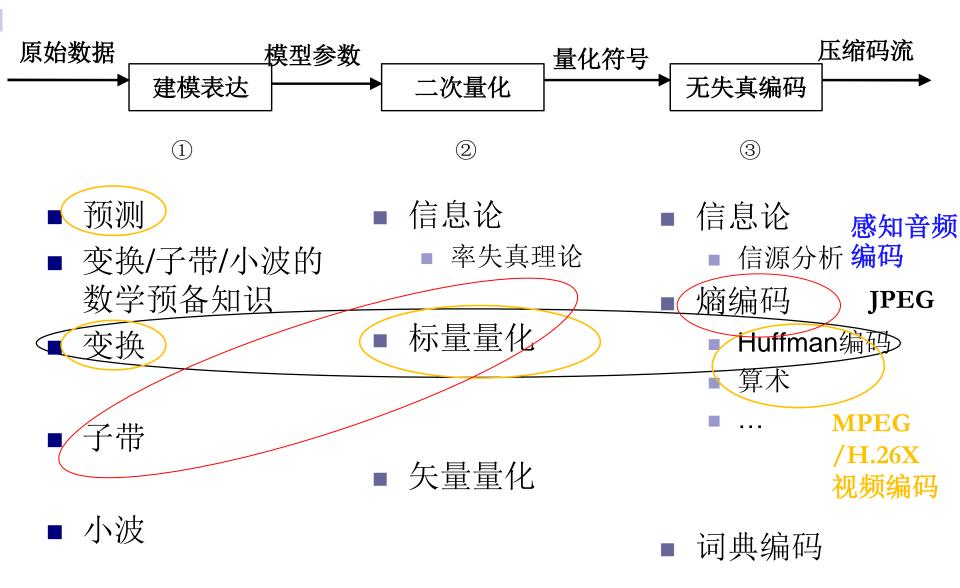
- ① 建立一个数学模型,以便更有效地重新表达原始数据
- ②对模型参数进行量化(二次量化,对模拟信号进行的 量化称为一次量化)
- ③ 更紧凑地表示量化后的模型参数,此时的编码要求能 "忠实地"再现模型参数的量化符号(熵编码)

#### 讨论:

如没有②且建模表示是一个可逆过程,则从压缩后的码流中可能完全恢复原始数据——可逆编码/无失真编码 几乎所有的有失真编码系统中都包括一个失真编码系统 **先讨论无失真编码缩,然后过渡到有失真编码** 

Introduction

#### 学习路线

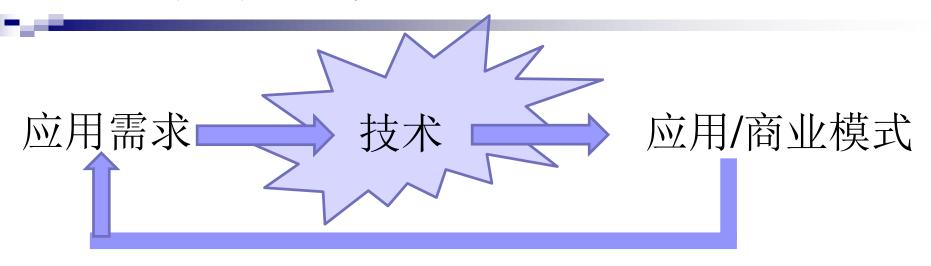


## 数据压缩技术的简单分类

		1					
熵编码	统计编码	霍夫曼编码、游程编码、二进制信源编码等					
		算术编码					
		字典编码: LZW编码等					
	其他编码	完	完全可逆的小波分解+统计编码等				
/ // // // // // // // // // // // // //	特征抽取	分析/综合		、 <i>ル</i> 白 <i>T</i> T	子带、小波、分形、模型基等		
	量化			î 狮 吗	其他		
		无记忆量化			均匀量化、Max量化、压扩量化		
		有记忆量化	序列量化	预测编码	增量调制、线性预测、非线性预测、自适应预测、运动补偿预测等		
				其他方法	序贯量化等		
			分组量化	直接映射	矢量量化、神经网络、方块截尾等		
				变换编码	正交变换、KLT、DCT、DFT、WHT等		
					非正交变换		
					其他函数变换等		

1-43

- 存储和传输的需求不同——流媒体?
  - > 数据压缩带来的挑战! 同步成为一个大问题!
- 渐进传输——提高用户体验质量?
  - 》有些编码算法允许先发送重要的的比特,使得 我们在看到更逼真的内容之前,能先看到更低 分辨率的版本
- 降低计算——移动设备的应用
  - > 用较少的计算得到近似的结果



- 同步问题的解决方案?
  - > 分解问题, 分层突破
- 渐进传输?
  - > 设计可分级的编码方案
- 降低计算?
  - > 大量的快速算法

- 数据压缩技术的发展及应用遵循"<u>标准先</u> <u>行</u>"的策略
  - ▶国际和国内技术标准的制定不仅前瞻性地预测了应用的需求,同时也为技术研究设定了明确的阶段目标。
  - ▶ 几十年来该领域的技术发展大大受益于国际标准化活动。

■ 与媒体表示和压缩有关的标准基本分为两类:

~以ISO为代表,ISO和IEC成立了ISO-IEC/JTC1(联合技术委员会),JTC1分为若干个子委员会(SC)

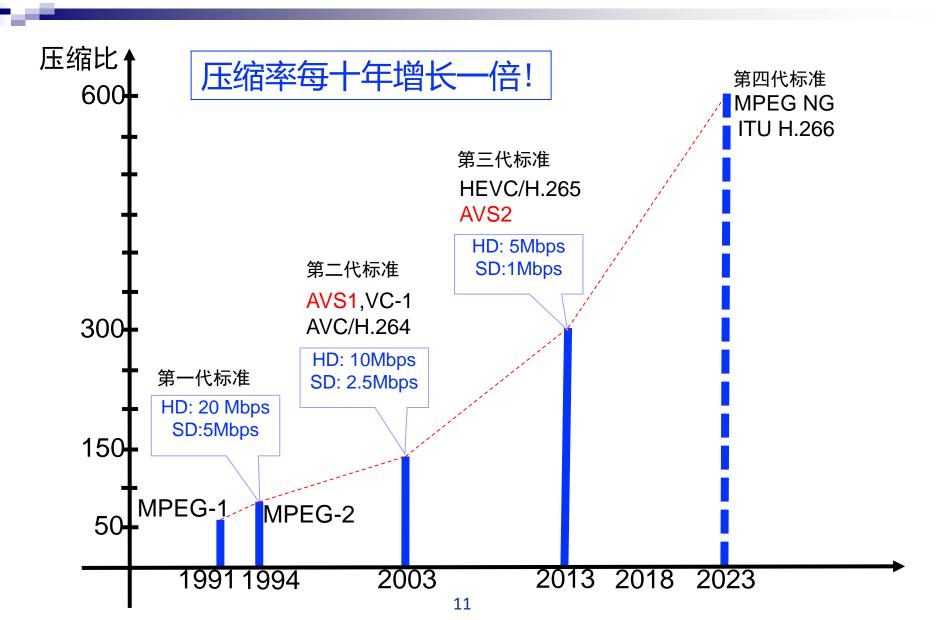
以计算机为中心,强调存储和回放。

www.mpeg.org

~以ITU为代表,成立了ICG/AVMMS(视听多媒体业务联合协调组)

以通信为基础,强调通信网络上多媒体信息的传输。

### 挑战之一: 编码上限?



### 主要的数据压缩编码标准

标准号	俗称	适用信源	典型应用
ITU-T T .82   ISO/IEC 11544	JBIG-1	二值图像、图形	G4传真机、计算机图形
ISO/IEC 14492	JBIG-2	二值图像、图形	传真、www图形库、PDA等
ITU-T T .81   ISO/IEC 10918	JPEG	连续色调静止图像	图像库、传真、彩色印刷、数码相机
ITU-T T .87   ISO/IEC 14495	JPEG-LS	连续色调静止图像	医学、遥感图像资料的无损/近似无损 压缩
ISO/IEC 15444	JPEG2000	连续色调静止图像	各种图形、图像(含计算机生成的)
ITU-T G.723、G.728和G.729		语音	数字通信和电话录音
ITU-T H.261建议	P×64	活动图像	ISDN上的会议电视/可视电话
ITU-T H.263建议		活动图像	PSTN上的会议电视/可视电话
ISO/IEC 11172	MPEG-1	活动图像及伴音	VCD、DAB、多媒体、VOD等
ITU-T H.262   ISO/IEC 13818-2	MPEG-2视频	高质量活动图像	SVCD/DVD、VOD/MOD、多媒体视 频游戏、DVB、DTV/HDTV等
ISO/IEC 13818-3	MPEG-2音频	高质量多声道声音	DAT、DCC、DAB等级数字视频伴音
ISO/IEC 14496	MPEG-4	多媒体音像数据	www上的视频、音频扩展
ITU-T H.264 (MPEG-4 part 10)	AVC	各种活动图像	H. 261/263和MPEG-1/2应用的替代
中国先进音、视频编码系列标准	AVS标准	活动图像和音频	广播电视、音像产品和多媒体通信等
ITU-T H.265	HEVC	各种活动图像	目前最新的视频编码标准

#### 1.4 编码系统的性能评价

- ■效率: 压缩比
- ■质量:对有失真编码而言
  - >客观质量: 失真度量
  - ▶主观质量
- ■复杂度
  - ▶算法的运算量和存储量
  - >编码器和解码器的复杂度:对称/非对称
- ■时延

### 压缩比

压缩比 = 
$$\frac{输入流的大小}{输出流的大小} = \frac{|x|}{|y|}$$

■ 压缩增益 = 100ln -

参考大小

压缩后的大小

- 无失真编码:
  - $\mathbf{x} = \hat{x}$  压缩比通常不超过1:4  $\mathbf{x} = \hat{x}$
- 有失真编码:
  - ightharpoonup 压缩比通常超过1:10  $x \neq \hat{x}$

输入流大小/某种无失真编 一码方法所产生的输出流

#### 失真的度量

■ 最常用:均方误差(Mean Squared Error, MSE)

$$e_k = x_k - \hat{x}_k$$
  $\sigma_e^2 = \mathbf{E}\left\{e_k^2\right\}$ 

■ 等价于信噪比(Signal Noise Ratio, SNR)

$$SNR(dB) = 10 \lg \frac{\sigma_x^2}{\sigma_e^2}$$

- 优点
  - > 易于计算
  - > 数学上在优化问题中易于处理
- ■缺点
  - > 没有考虑人的主观感知特性

#### 失真的度量:对图像

■均方误差:

$$MSE = \frac{1}{M \times N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \left\{ x(m,n) - \hat{x}(m,n) \right\}^{2}$$

■峰值信噪比( Peak Signal Noise Ratio, PSNR )

$$PSNR(dB) = 10 \lg \frac{x_{\text{max}}^2}{\sigma_e^2}$$

#### 问题分析: 客观指标与主观感知的差异?





峰值信噪比=5.98

峰值信噪比=6.24

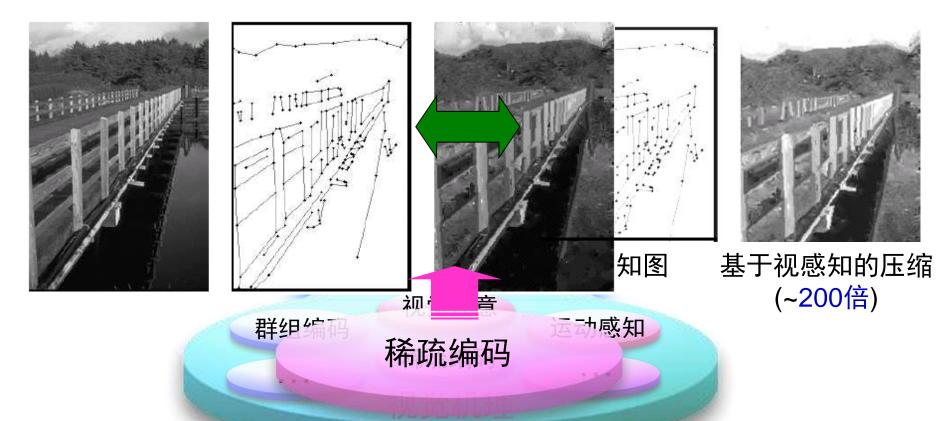
一般认为:峰值信噪比与图像质量近似成正比关系。

结论: 1、峰值信噪比度量与人的视觉感知并不完全一致!

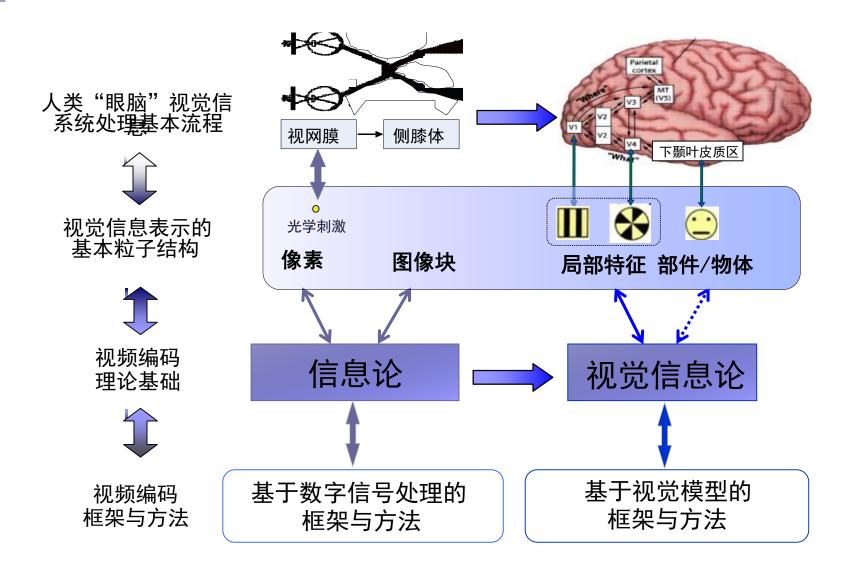
2、需要寻求更加符合人类视觉感知的客观度量方法以 及相应的编码理论和方法。

#### 突破方向: 借鉴视觉感知机理

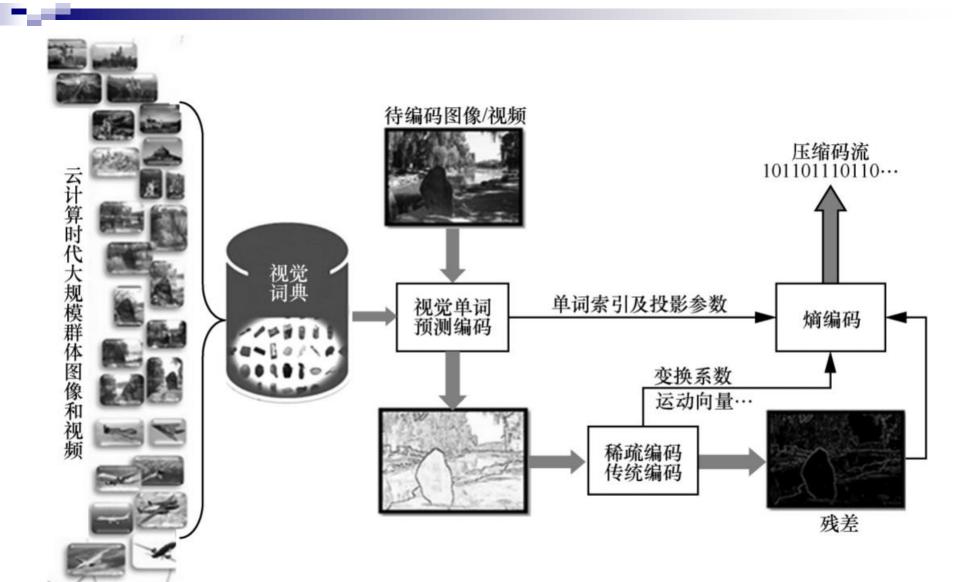
❖对视觉感知机理的认识逐渐深入,用于视频编码的潜力很大



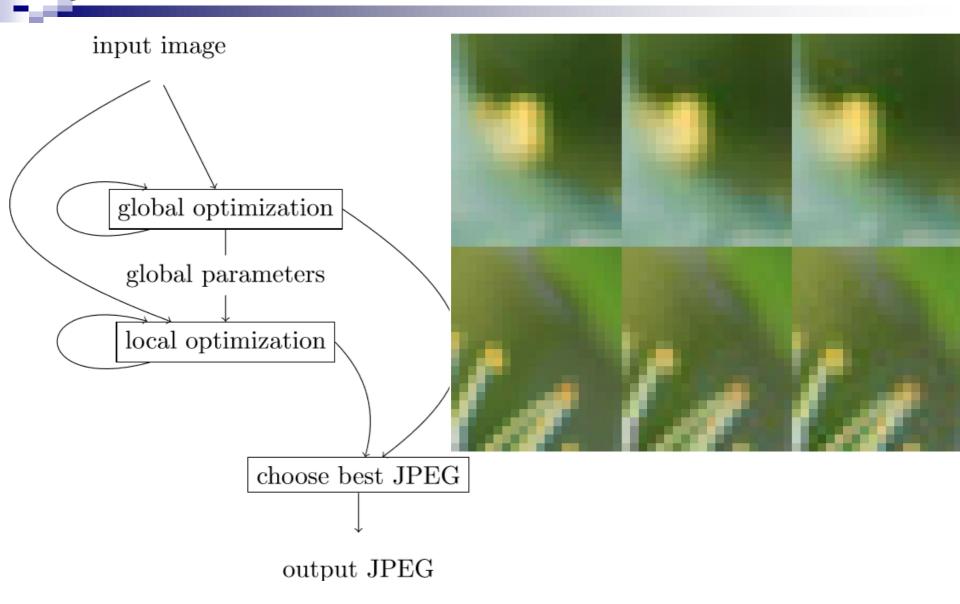
#### 基于视觉特性的视频编码



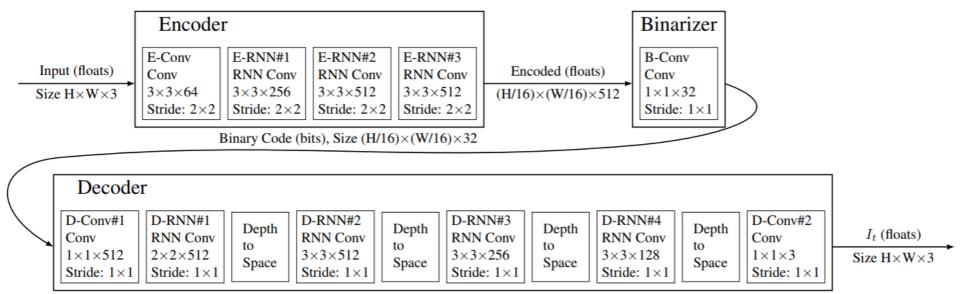
## 云编码



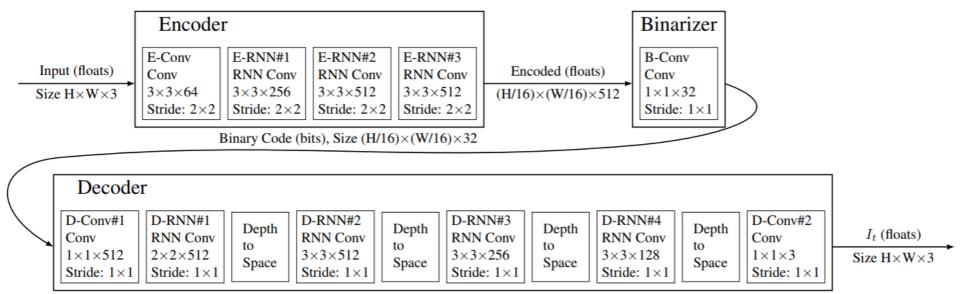
## Google Guetzli- Perceptually Guided JPEG Encoder 2017



# MIT-基于递归神经网络的全分辨率图像压缩 CVPR2017



# MIT-基于递归神经网络的全分辨率图像压缩 CVPR2017



#### 1.5 本课程内容

- 第一部分: 视音频媒体的特性与表示
  - ▶图像的属性和常见文件格式,视频信号的数字 化及相关标准,AVI等常见的视频文件格式, 音频的特性和常见文件格式
- 第二部分: 无失真压缩的理论基础
  - >信息论基础: 信源的建模与分析
  - > 变长码的编码基础
    - 无失真信源编码定理
  - ▶编码的基本途径

#### 1.5 课程内容

- 第三部分: 统计编码方法
  - > Huffman编码,Golomb编码和通用变长码, 算术编码,游程编码
  - ▶词典编码(LZ77、LZ78、LZW算法,词典编码在文件压缩和图像压缩中的应用)
- 第四部分: 有失真压缩的理论基础
  - > 失真的度量,率失真理论简介
  - ▶ 有失真压缩中常用的模型: 概率模型和线性系 统模型;

#### 1.5 课程内容 (续)

- 第五部分: 有失真信源编码(熵压缩)
  - ▶量化
  - ▶基于变换的编码、JPEG标准
  - > 预测编码
  - ▶\*子带编码、小波编码、 JPEG2000
- 第六部分: 信源编码标准及应用
  - ➤ 视频编码标准 (MPEG-2, MPEG-4, H.264, AVS,...)
  - ▶音频编码标准(MP2, MP3, G.721, G.729,...)
  - ➤应用
    - ■码率控制、转码、传输

#### 本课程的学习内容

- 只有站在一定的高度来观察整个系统,才可能比较深刻地理解每个(工具)模块的作用和整体优化方法,这是自顶向下去理解问题的方式。
- 学习每种数据压缩工具(算法)的原理及实现方法,才有可能在面对具体的数据压缩问题时选择适当的工具并加以组合。
- 数据压缩是一门科学,更是一门**实验科学**, 要对一种算法有深入的理解,最好的方法 就是实现(至少调试)该算法。
- 实验非常重要!

#### 本课程的目的

- ■掌握数据压缩的基本理论、标准及应用
  - 》掌握视音频媒体的特性和文件格式分析方法
  - 无失真压缩和有失真压缩技术的基本原理及相 关应用
  - > 培养解决问题的思维方法,能综合运用于对实际问题的分析
  - 〉初步具有应用所学技术解决实际问题的能力
  - > 培养工科学生的综合工程素质