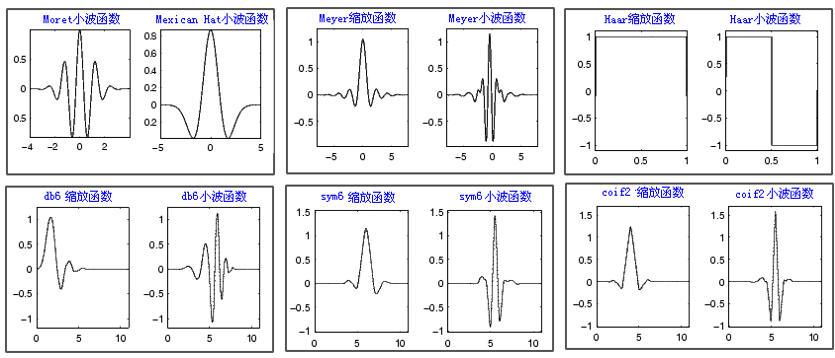
**第六章小波与小波变换**

1、**小波概念**

在有限时间范围内变化且其平均值为零的数学函数 具有有限的持续时间和突变的频率和振幅 在有限的时间范围内，它的平均值等于零



2、**小波变换**

----小波变换(wavelet transform，WT)是用小波对函数在空间和时间上进行局部化分析的数学变换

\*\*小波变换通过平移母小波获得信号的时间(位置)信息，通过缩放母小波的宽度(或称尺度)获得信号的频率特性。对母小波的缩放和平移是为了计算小波的系数

\*\*傅里叶变换提供信号包含的频率信息，但时间方面的局部化信息基本丢失。小波变换继承和发展了傅里叶变换、哈尔变换和短时傅里叶变换(STFT)的思想，不仅可提供信号包含的频率，而且还可提供信号的时间和频率之间的关系。

**连续小波变换：**连续小波是指没有经过数字化的小波 ，连续小波变换(CWT)是用连续小波表示函数的数学变换，用于对连续函数在时间和空间上进行局部化分析。理解CWT可从傅里叶级数和傅里叶变换开始。

\*\***小波变换的概念** 傅里叶分析把一个信号分解成频率不同的正弦波，正弦波是傅里叶变换的基函数。小波分析把一个信号分解成一系列小波之和，这些小波是母小波经过移位和缩放之后的小波，同样可以用作表示函数的基函数。 凡能用傅里叶分析的函数都可用小波分析，小波变换可理解为用一系列小波代替傅里叶变换中的正弦波。 用不规则的小波分析变化激烈的信号，比用平滑的正弦波更有效，对信号的基本特性描述得更准确。

----**傅里叶变换** 傅里叶级数是用简单的三角函数之和来表示函数的方法。简单是指三角函数可用振幅、频率和相位三个要素表示即可 可把任意的周期函数或周期信号分解成一系列正弦和余弦函数之和。例如，方波函数 可用正弦波函数之和来表示。一个函数可用频率不同的正弦波之和表示，这些正弦波称为傅里叶变换的**基函数**(basis function)。**基函数是函数空间中的基本元素**，函数空间中的每个连续函数都可表示为基函数的线性组合

数学上，连续小波变化公式为：，其中(1) 小波变换是信号f(t) 与缩放和平移的小波函数ψ之积在信号存在的整个期间里求和。(2) 连续小波变换(CWT)的结果是小波的系数C，这些系数是缩放因子(scale)和位置(position)的函数。如用a表示缩放因子，对于正弦函数，f(t)=sint=>a=1,f(t)=sin2t=>a=1/2,f(t)=sin4t=>a=1/4。

**小波与傅里叶的区别：**小波与傅里叶变换是多媒体技术中常用的信号分析工具，它们之间有一些重要的区别：

\*\*\*时间-频率分辨率：小波变换可以提供随时间变化的频率信息，因此在处理非平稳信号（随时间变化的信号）时具有优势。而傅里叶变换则提供信号在整个时间段内的频率信息，对于平稳信号有更好的适用性。

\*\*\*层级分解：小波变换可以将信号进行多尺度的分解，从整体到局部进行分析，因此能够更好地捕捉信号的局部特征。而傅里叶变换是在整个时间范围内对信号进行频域分析，无法提供多尺度分析的信息。

\*\*\*时间域信息保留：小波变换在进行信号分解时可以保留更多的时间域信息，这对于一些需要保留时域特征的应用有重要意义。而傅里叶变换将信号完全转换到频域，丢失了时间域信息。

\*\*总的来说，小波变换适用于非平稳信号和需要局部特征分析的场景，而傅里叶变换更适合于平稳信号和全局频率分析的场景。在实际应用中，两者经常结合使用，根据需要选择合适的方法进行信号分析。

**小波变换步骤，**可分如下5步 ：

---1小波ψ (t)和原始信号f(t)的开始部分进行比较

---2计算系数C——该部分信号与小波的近似程度；C值越高表示信号与小波相似程度越高，因此系数C可以反映这种波形的相关程度

---3小波右移距离k得到的小波函数为ψ (t-k) ，然后重复步骤1和2，再把小波向右移，得到小波ψ (t-2k)，重复步骤1和2……直到信号f(t)结束

---4扩展小波，如扩展一倍，得到的小波函数为ψ (t/2) ----5重复步骤1~4

**小波基函数**：一个函数可用小波基函数的线性组合表示。小波：由定义在有限区间里具有基本形状的函数ψ(x)构造的小波。母小波(mother wavelet)：构造小波的函数ψ(x)。子小波(daughter wavelet)：母小波函数ψ(x)构造的小波。小波基函数(wavelet basis function)：通过缩放和平移母小波ψ(x)构造的一系列小波函数

**离散小波变换**：使用离散小波的小波变换。离散小波：连续小波通过采样和量化后的小波。小波变换计算：连续小波变换：实际用离散数据计算，缩放因子和平移参数比较小。离散小波变换：缩放因子和平移参数均选择 2^i(i >0的整数)，称为二进小波变换(dyadic wavelet transform)。离散小波分析得到的时-频(时间-频率)关系。**执行离散小波变换的有效方法是使用滤波器。小波分解树**：由低通滤波器和高通滤波器组成分解树。**小波包分解树：**对高频分量和低频分量连续分解，小波包分解方法是小波分解的一般化。**注意：使用滤波器对真实的数字信号进行变换时，得到的数据将是原始数据的两倍。**

**小波重构**：把分解的系数还原成原始信号的过程称为小波重构或合成，在数学上称为逆离散小波变换。**两个过程**：在小波重构时包含升采样和滤波两个过程。

3、**哈尔小波变换**

**哈尔小波基函数**：基函数是一组线性无关的函数，可以用来构造任意给定的信号。哈尔小波是小波中最简单的小波。哈尔小波定义为：[0,1/2)上f(x)=1,[1/2,1)上f(x)=-1,其他f(x)=0。f(x)也称为母小波。

**一维哈尔小波变换**：小波变换的基本思想是用一组小波函数表示一个函数或信号。一维离散哈尔小波变换是用一维离散哈尔小波表示一维离散数据

【例】 假设只有4个像素的一维图像，对应图像位置的像素值分别为 =[9, 7, 3, 5]，即x1=9,x2=7,x3=3,x4=5，计算一维离散哈尔小波变换的系数。

步骤1：求均值(averaging) \*\*平均值：相邻一对像素的平均值 \*第一对像素的平均值为(9+7)/2=8 \*第二对像素的平均值为(3+5)/2=4 \*\*存储像素的平均值[8, 4] \*\*变换后的图像分辨率为原图像的1/2

步骤2：求差值(differencing) \*\*差值：相邻一对像素的差值的平均值 \*第一对像素的细节系数：(9－7)/2=1 \*第二对像素的细节系数：(3－5)/2 =－1 \*\*存储像素的细节系数[1, -1] 原始图像数据[9, 7, 3, 5]变换成[8, 4, 1, -1]

步骤3：重复步骤1和2 把由第一步分解得到的图像数据[8, 4, 1, -1]进一步分解，得到分辨率更低的图像[6，2]，分解过程见表6-1

**哈尔小波变换的结果** 把原始图像的数据[9, 7, 3, 5]用[6, 2, 1, -1]表示，即由4个像素组成的图像用1个平均值和3个细节系数表示 这个过程叫做哈尔小波变换(Haar wavelet transform)，也称哈尔小波分解(Haar wavelet decomposition) 这个概念可以推广到使用其他小波的小波变换。

**规范化的哈尔小波变换**与非规范化的哈尔小波变换相比，唯一的差别是在规范化的变换中用21/2 代替2，这是因为要满足小波基函数为正交基的要求，但并影响对哈尔小波变换的理解。

**二维小波变换**：图像块是二维数据阵列，小波变换可对每一行做变换，然后对行变换后的每一列做变换。用小波对图像进行变换有两种方法：标准分解(standard decomposition)，非标准分解(nonstandard decomposition)。两种方法比较：变换结果完全相同，非标准分解算法的计算量较少

**第七章 小波图像编码**

**子带编码**：把信号频率分成几个子带，对每个子带分别进行编码。根据每个子带的重要性分配不同位数表示数据。

**小波图像编码**：基于小波的图像编码。

1、小波图像分解

**小波分解图像方法：**均匀分解法(uniform decomposition) 非均匀分解法(non-uniform decomposition) 八带分解法(octave-band decomposition) 小波包分解法(wavelet-packet decomposition) 自适应小波分解法(adaptive wavelet decomposition)。**八带分解是使用最广泛的分解方法，**属于非均匀频带分割方法，把低频部分分解成比较窄的频带，对每一级分解的高频部分不再进一步分解。E.g将图像一级分解为近似值，水平垂直细节，对角细节，二级分解只分解近似值那块。

**评估编码系统性能的方法**：失真度量法——用峰值信号与噪声之比衡量, 定义为最大信号峰值的平方与信号的均方差之比；规格化均方差；信号噪声比和平均绝对误差。

**EZW编码（嵌入型零树小波编码）**：包含三个概念。小波：指该算法以离散小波变换为基础；零树：小波变换系数之间的数据结构；嵌入：渐进编码技术的另一种说法。主要由小波变换、量化和熵编码三个模块组成。是多分辨率图像的一种编码方法。

**EBCOT编码（最佳截断嵌入码块编码）**：基本思想：把每个子带的小波变换系数分成独立编码的码块(code-block)，对所有码块使用完全相同的编码算法。

2、**JPEG、JPEG2000压缩算法**

(1)**JPEG是什么** \*\*\*JPEG=Joint Photographic Experts Group \*\*静态数字图像数据压缩标准 \*由ISO和IEC两个组织联合组成的专家组开发 \*国际通用标准，称为JPEG标准 \*采用JPEG标准压缩的文件使用.JPG或.JFF作为文件扩展名\*\*\* JPEG专家组开发了两种基本的压缩算法 \*\*采用以离散余弦变换(Discrete Cosine Transform，DCT)为基础的有损压缩算法。压缩比为25:1时，非图像专家难以找出压缩前后图像之间的的差别 \*\*采用预测技术为基础的无损压缩算法 \*\*\*JPEG专家制定了采用小波变换的JPEG 2000标准

(2)JPEG算法概要 利用视觉系统特性，使用变换、量化和熵编码相结合，去掉或减少视觉冗余和数据冗余 压缩算法大致分成三个步骤 用正向离散余弦变换(forward discrete cosine transform，FDCT)把空间域表示的图变换成频率域表示的图 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 使用霍夫曼可变字长编码器对量化系数进行编码 颜色空间变换(RGB到YUV)和(YUV到RGB)不包含在JPEG算法，处理单独的彩色分量图像

(3)JPEG算法的计算步骤 (1) 正向离散余弦变换(FDCT) (2) 量化(quantization) 对经过FDCT变换后的频率系数进行量化。 目的：降低非0系数的幅度，增加0值系数的数目。 量化是造成图像质量下降的最主要原因。 使用如图4所示的均匀量化器进行量化 量化步距按系数所在位置和每种颜色分量的色调值确定。使用两种量化表 表5-6用于亮度量化表，表5-7用于色度量化表 人眼对低频分量图像比高频分量图像更敏感，因此表中的左上角的量化步距要比右下角的量化步距小 表中数值对CCIR 601标准电视图像是最佳的(3) Z字形编码(zigzag scan) 3. Z字形编排 重新编排量化后的系数，增加连续的0值系数数目 排列方法：按Z字形排列，如图5(a) DCT系数序号见图5 (b)，序号小的位置表示频率较低，用 zz(0)，zz(1)，…，zz(63)表示 8×8的矩阵变成1× 64的矢量(4) 使用DPCM对直流系数(DC)进行编码 (5) 用行程长度编码(RLE)对交流系数(AC)进行编码 (6) 熵编码(entropy coding)目的：压缩采用DPCM编码后的DC系数差值和RLE编码后的AC系数 方法：使用霍夫曼编码器，用查表方法编码，霍夫曼码表可事先定义（7） 组成数据位流 JPEG编码的最后一个步骤，把各种标记代码和编码后的图像数据组成一帧一帧的数据，便于传输、存储和译码器译码

**JPEG2000**

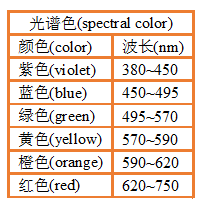
JPEG 2000是以小波技术为基础的静态图像压缩编码标准。JPEG 2000是JPEG委员会在2000年确立的标准,标准号为ISO/IEC 15444,文件扩展名为.jp2(用于ISO 15444-1)和.jpx(用于ISO15444-2)。**JPEG 2000目标**是提高对连续色调图像的压缩效率,管理和传输,而又不使图像质量有明显的损失。该标准使用小波技术以提高压缩比,用户可控制图像的分辨率,在网络上传输时可按照用户要求下载各种分辨率的图像。此外,该标准可提供无损压缩的图像,在文档中可提供更多的颜色信息。

**编码器基本结构：**源图像小波变换⇒系数量化⇒熵编码⇒代码流/位流 **解码器基本结构：**码流熵解码⇒逆量化⇒逆向变换⇒重构图像

--**编码方法：**基于EBCOT算法，以图像块为处理单元，块大小无限制。图像变换、量化和熵编码等均以图像块为单元 **优点**：可降低对存储器的要求，便于抽出图像中的部分图像 **缺点**：图像质量有所下降，但不明显

---**主要特性** 压缩率比JPEG高20%以上，质量相同。渐进传输：可先传输低分辨率图像或其轮廓，然后逐步传输其他数据，不断提高图像质量。兴趣区编码：可对感兴趣的图像区单独编码，指定特定的质量要求。JEPG2000的标准文档 包括14个部分(Part)， 最核心文档 Part 1：Core coding system(核心编码系统)

**第八章 颜色度量体系**

**颜色是什么**：视觉系统对可见光的感知结果，感知到的颜色由光波波长决定。波长范围：380～780 nm。纯颜色用光的波长定义，称为光谱色(spectral colors)。用不同波长的光进行组合可产生相同的颜色感觉。区分颜色的**三个特性：色调；饱和度；明度**

**色调**：视觉系统对一个区域呈现的颜色的感觉，即对可见物体辐射或发射的光波波长的感觉。用色调最容易把颜色区分开。色调用红(red )、橙(orange)、黄(yellow)、绿(green)、青(cyan)、蓝(blue)、靛(indigo)、紫(violet) 等术语刻画。用于描述感知色调的术语是色彩(colorfulness)，如浅蓝或深蓝的感觉。黑、灰、白为无色彩

色调在颜色圆上**用圆周表示**。圆周上的颜色具有相同饱和度和明度，但色调不同。色调数目多于1000万种

普通人可区分200种色调、50种饱和度和500级灰度。颜色专业人士可辨认的色调数大约300～400种。

**饱和度**：颜色的纯洁性。可用来区别颜色明暗的程度；一种颜色掺入其他光成分越多，颜色越不饱和；完全饱和的颜色是指没有渗入白光所呈现的颜色；单一波长的光谱色是完全饱和的颜色。**半径表示法**：沿径向方向上的饱和度不同，色调和明度相同，七种颜色的饱和度不同，左边的饱和度最浅，右边的饱和度最深 ，它们色调和明度相同

**明度**：视觉系统对物体辐射光或发射光多少的感知属性。有色表面的明度取决于亮度和表面的反射率。感知的明度与反射率不成正比，认为是一种对数关系。明度的主观感觉值目前无法用物理设备测量。可用亮度(luminance)即辐射的能量来度量，用一个数值范围表示，如 0～10。黑色(没有光)和白色是两个极端，在它们之间的颜色是灰色。**垂直轴表示法：**7种颜色具有，相同色调和饱和度，不同的明度，底部的明度最小，顶部的明度最大。

**亮度**：国际照明委员会定义的物理量。用辐射功率度量。用反映视觉特性的光谱敏感函数加权后得到，在555 nm处达到峰值，幅度与物理功率成正比。亮度犹如光的强度(intensity)。在CIE XYZ系统中，亮度用Y表示，用单位面积上反射或发射的光强度来度量。明度和亮度的关系不是线性关系，也不是同义词。亮度应使用烛光/平方米(cd/m2)作度量单位，但实际上用指定的亮度(白光)作参考，并把它标称化为1或者100个单位。例如，监视器用亮度为80 cd/m2的白光作参考，并指定Y=1

**光亮度**：根据CIE的定义，光亮度(lightness)是人的视觉系统对亮度(luminance)的感知响应值，并用L\*表示。与明度相同，光亮度也常作为颜色空间的一个维。

**颜色空间**：也叫**颜色模型**，是一种用空间中的点表示颜色的数学表示法。**对人**，可通过色调、饱和度和明度定义颜色；**对显示设备**，可用红、绿和蓝磷光体的发光量描述颜色；**对打印或印刷设备**，可用青色、品红色、黄色和黑色的反射和吸收来产生指定的颜色。**通常用三维模型表示**。

**颜色空间类型**：设备相关——指定颜色与颜色生成设备有关，设备无关【例】RGB颜色空间是与显示设备相关的颜色空间。显示器用RGB显示颜色，如用R＝250, G＝123, B＝23生成的颜色，将随显示器的亮度和对比度改变——指定颜色与颜色生成设备无关。【例】 CIE L\*a\*b\*颜色空间是设备无关的颜色空间，它建筑在HSV(hue, saturation and value)颜色空间的基础上，用该空间指定的颜色无论在什么设备上生成的颜色都相同。

**颜色系统**：颜色系统=颜色度量体系=颜色制，是一种组织和表示颜色的方法。**曼塞尔系统**(Munsell system)：用色调（RYGBP五个及5个相邻交叉色调）、明度（11个等级）和色度（15个等级）定义和排列颜色；**Ostwald系统**(Ostwald system)：用波长、纯度和亮度映射色调、饱和度和明度；**CIE颜色系统**(CIE color system)：1931年首次用数学方法定义颜色。

**组织和表示颜色的两种方法**：颜色模型：用数值方法指定颜色e.g RGB, HSB, CMY, CIE XYZ, CIELAB和用光谱描述。编目系统：为每种颜色分配唯一的名称或号码。E.g Munsell颜色系统，Pantone颜色系统。

**物理三基色模型**：RGB模型采用物理三基色，其物理意义很清楚，但它是一种设备相关的颜色模型。**理论三基色模型**：为了解决这个问题，国际照明委员会的颜色科学家们企图在RGB模型基础上，用数学方法从真实的基色推导出理论的三基色，创建一个新的颜色系统，使颜料、染料和印刷等工业能够明确指定产品的颜色。

**CIE 1931 RGB系统**：

**1.RGB颜色匹配函数**：按照三基色原理，颜色实际上是物理量，对物理量可进行计算和度量。根据这个原理产生了用红、绿和蓝单光谱基色匹配所有可见颜色的想法。横坐标：光谱波长，纵坐标：匹配光谱各色所需要的RGB数值。**CIE 把三种单色光的波长分别定为**：红光：700 nm；绿光：546.1 nm；蓝光：435.8 nm。

**2.标准白光Ew的匹配**：R=1个红基色光单位=1光瓦；G=1个绿基色光单位=4.5907光瓦；B=1个蓝基色光单位=0.0601光瓦，其中，1光瓦=680流明(lm)。则标准白光Ew的颜色用每个基色单位为1的物理三基色配出——C\_Ew=1\*R+1\*G+1\*B。其他颜色则用C=rR+gG+bB配出，rgb分别是三基色的比例系数。r、g和b的大小决定颜色光的光通量。r、g和b之间的比例决定生成的颜色。因为三基色的总光通量必须与被表示的颜色相等，因此 r、g和b之和必等于1，也就是r+g+b=1。

**CIE XYZ系统**：设备无关的颜色系统。根据视觉特性和用颜色匹配实验结果定义。规定X、Y和Z基色都用正数去匹配所有的颜色，并用Y值表示人眼对亮度(luminance)的响应。每种颜色都可表示成想象的基色X、Y和Z的混合。与可见的颜色不对应。相较于CIE RGB系统，颜色匹配函数纵坐标修改成了XYZ基色的相对值。采用XYZ基色进行组合颜色。缺点：使用比较复杂，且不直观。



**RGB和XYZ之间的转换关系**:公式如图。



**CIE xyY色度图**：从XYZ直接导出的颜色空间，使用亮度Y参数和颜色坐标x，y刻画颜色，Y值：颜色亮度或称光亮度，同XYZ中的Y刺激值，坐标x，y：指定二维色度图上的颜色。该色度图上可以直观的看出来色调和饱和度。等色调线就是从白光E点出发的径线，等饱和度线就是内部的闭合曲线。也可以看出来色域——闭合曲线所包围的区域。**彩色显像管用混色法重现颜色,不能直接采用CIE标准**。

**CIE YUV**:由CIE 1931 xyY经过线性变换之后得到的颜色空间。在色度图中，代表两种颜色的两个点之间的色差与对颜色感知的差别是均匀的。其中的Y与XYZ或xyY中的Y相同,u=2x/(6y-x+1.5),v=3y/(6y-x+1.5)。**CIE YU’V’**:进一步减少色差与感知的非线性。Y不变，u’=u=2x/(6y-x+1.5),v’=1.5v=4.5y/(6y-x+1.5)。**CIE Luv**：采用了UCS方案,先把亮度和颜色完全分开，然后使用数学公式把CIE 1931 XYZ中的x，y坐标变换到为u, v的新坐标系，得到比较精确的色度图。**CIE Lu’v’**:把u, v重新命名为u‘, v’ ,经过数学变换得到更加均匀的色度图。**CIE 1976 L\*u\*v\***:为解决刺激值Y(亮度)的非线性问题，CIE定义了一个均匀的光亮度比例，L\*，其值从0(黑色)～100(白色)。使用L\*，u’, v’定义了L\*u\*v\*颜色空间，称为CIELUV或Luv颜色空间，也称CIE 1976 L\*u\*v\*颜色空间。**CIE LAB**：度量颜色的国际标准，使用最广泛的物体颜色度量方法。可简写为CIELAB，也称CIE 1976 L\*a\*b\* (简写为CIE L\*a\*b\*)颜色空间，或称为CIELAB/CIEL\*A\*B\*色差制。特点：光亮度、色调和饱和度都可独立调整；与设备无关，可生成一致的颜色；使用对色 (opponent color)坐标。**CIELUV L\*C\*h\***是从CIE 1976 L\*u\*v\*导出的颜色空间.**CIELAB L\*C\*h\***是从CIE 1976 L\*a\*b\*导出的颜色空间。其中LCH分别是光亮度、色调和色度。

**第九章 颜色空间转换**

**颜色空间的分类**:从颜色感知的角度:1.从颜色感知的角度:混合(mixture)型：按三种基色的比例合成颜色.e.g RGB，CMY(K)和XYZ。非线性亮度/色度(luma/chroma)型：用一个分量表示非色彩的感知，用两个独立的分量表示色彩的感知，如L\*a\*b, L\*u\*v，YUV和YIQ。强度/饱和度/色调(intensity/saturation/hue)型：用饱和度和色调描述色彩的感知，颜色解释直观。如HSI, HSL, HSV和LCh。2.从应用角度区分：显示和打印（RGB,CMY），计算机图形(HSV,HSL,HSB,HVC)，设备无关(CIE系列)，电视系统(YUV,YIQ)。

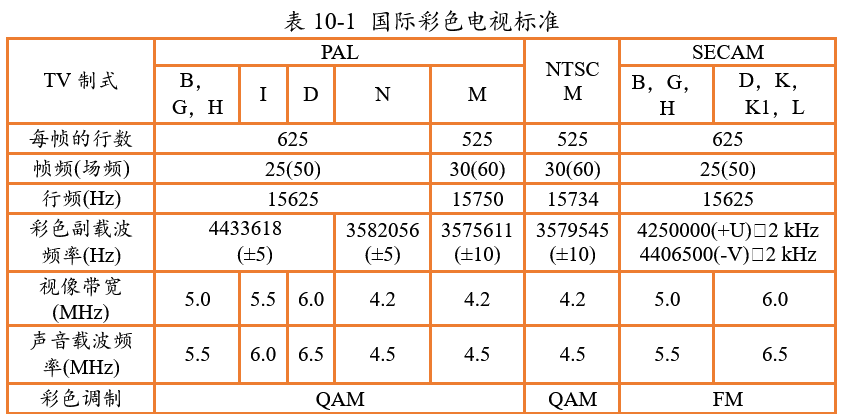
**RGB和CMY的转换公式**：C=1-R,M=1-G,Y=1-B,RGB的转化同理。**CMY->CMYK**: black=min(C,M,Y),C=(C-BLACK)/(1-BLACK),M=(M-BLACK)/(1-BLACK),

Y=(Y-BLACK)/(1-BLACK). **CMYK->CMY**:C=min(1,C\*(1-black)+black),M=min(1,M\*(1-black)+black),Y=min(1,K\*(1-black)+black)。

**电视系统的颜色空间**：颜色空间特点：亮度和色度相互独立，与设备相关。European Y'U'V'；American Y'I'Q'；SMPTE-C RGB；ITU-R BT.601 Y'CbCr；ITU-R BT.709 Y'CbCr；SMPTE-240M Y'PbPr；YCgCo颜色空间。

**第十章 数字电视基础**

**电视**：捕获、广播和重现活动图像和声音的远程通信系统。**电视系统类型**：黑白电视和彩色电视。**电视制**：传输图像和声音的方法。

**国际彩色电视标准**：如图。

**电视的扫描方式**:隔行扫描与逐行扫描。差别：**黑白电视和彩色电视都用隔行扫描.计算机显示图像时一般都采用逐行扫描。**隔行扫描中一帧图像由两部分组成：一部分是由奇数行组成，称奇数场，另一部分是由偶数行组成，称为偶数场，两场合起来组成一帧。在隔行扫描中，无论是摄像机还是显示器，获取或显示一幅图像都要扫描两遍才得到一幅完整的图像。

**电视扫描术语**：**f\_f：场频/场速率**，每秒钟扫描的场数。根据视觉特性和电网频率(50Hz或60Hz)确定，使在屏幕上显示的图像看起来不会让人感觉到在闪烁，以及减低电网频率的干扰。 **f\_F：帧频/帧速率**，每秒扫描的帧数。单位：帧每秒(frames per second，fps)。PAL制和NTSC制的帧频分别为25 fps和30 fps。**f\_H ：行频/水平行速率** , 每秒钟扫描的行数。【例】NTSC制精确的帧频是210.97 Hz，525行每帧，因此行频为210.97×525＝15 734行/秒

**彩色电视制**：**PAL制的扫描特性**：一帧图像的总行数为625，分两场扫描，场频为50 Hz，周期为20 ms，帧频为25 Hz，周期为40 ms ，行扫描频率(行频)为15 625 Hz，周期为64μs，每一行传送图像的时间为52.2μs，行扫描逆程时间为11.8μs ，兼作行同步和消隐。黑白电视的扫描定时与同步信号，包括水平同步(即行同步)、场同步等信号。**NTSC制的扫描特性**：一帧图像的总行数为525行，分两场扫描。场扫描频率是60 Hz，周期为16.67 ms。每场扫描行数：525/2=262.5行。每场开始保留20条线作为控制信号，一帧485条线可见。 帧频(刷新频率)30 Hz(精确29.97)，周期33.33 ms。行扫描频率为15 750 Hz，周期为63.5μs，其中水平回扫时间为10μs(包含5μs的水平同步脉冲)，显示图像的时间为53.5μs。**SECAM制的扫描特性**：和PAL一样。

**三种类型的彩色电视信号**：复合电视信号，分量电视信号，S-Video信号。

**图像子采样概要**：**两种采样方法：**对亮度信号和色差信号采用相同的采样频率采样；对亮度信号和色差信号采用不同的采样频率采样 **图像子采样：**色差信号的采样频率比亮度信号的采样频率低。最简单的视像压缩技术，依据是视觉系统的特性。对色度的敏感比对亮度的敏感低，把人眼刚能分辨的黑白相间条纹换成不同颜色的彩色条纹，眼睛不再能分辨出单独的条纹，可去掉部分颜色数据而不易察觉；对图像细节的分辨能力有一定的限度，可去掉图像中的高频信号而不易察觉

**常见的采样格式**

\*\*\*4:4:4 这种采样格式不是子采样格式，它是指在每条扫描线上每4个连续的采样点取4个亮度Y样本、4个红色差Cr样本和4个蓝色差Cb样本，每个像素用3个样本表示

\*\*\*4:2:2 在每条扫描线上，每4个连续的采样点取4个亮度Y样本、2个红色差Cr样本和2个蓝色差Cb样本，平均每个像素用2个样本表示

\*\*\*4:1:1 在每条扫描线上，每4个连续的采样点取4个亮度Y样本、1个红色差Cr样本和1个蓝色差Cb样本，平均每个像素用1.5个样本表示

\*\*\*4:2:0 在水平和垂直方向上，每2个连续采样点上取2个亮度Y样本、1个红色差Cr样本和1个蓝色差Cb样本，每个像素用1.5个样本表示

**4: 4: 4 YCbCr格式：**对每个采样点，Y, Cb, Cr各取一个样本。消费类和计算机应用：每个分量的样本精度为8位。编辑类应用：每个分量的样本精度为10位。每个像素的样本需要24位或30位。**4: 2: 2 YCbCr格式**：每2个Y样本有1个Cb和1个Cr样本。消费类和计算机应用：每个分量的样本精度为8位。编辑类应用：每个分量的样本精度为10位。仅Y样本的Cr和Cb ：用前后相邻的Cr和Cb样本计算。**4: 1: 1 YCbCr格式**：每4个Y样本各有1个Cb和1个Cr样本，每个分量的样本精度为8位。仅Y样本的Cr和Cb ：用前后相邻的Cr和Cb样本计算。应用：数字电视盒式磁带**4: 2: 0 YCbCr格式**：在水平和垂直方向上各有2个Y样本，用4个样本点的色差计算得到1个Cb和1个Cr样本，合计6个样本。每个分量的样本精度为8位或10位。 **MPEG-2和MPEG-1都用4:2:0格式**，但MPEG-2在水平方向上没有半个像素的偏移。

**数字电视**：DTV = digital television / digital TV。**定义1**: 用数据压缩和数字传输技术传送视像和声音的广播通信系统。**定义2**：用数字形式表示的活动图像和声音。**传输方式**：**卫星转播**：卫星数字电视(digital satellite TV) **地面广播**：地面数字电视(digital terrestrial TV) **电缆传输**：有线数字电视(digital cable TV)

**第十一，十二章MPEG**

**MPEG的重要性**：VCD影视是MPEG-1标准的典型应用。MP3是MPEG-1标准另一个大规模应用的实例。MPEG-1, -2已成为数字电视设备的心脏，如VCD/DVD播放器和数字电视机顶盒等。MPEG-1=>数字电视，VCD影视，MPEG-2=>数字电视，DVD影视。

1.**MPEG-1数字电视标准**

(1)**MPEG-1是什么**：1992年发布的第一个数字电视编码标准，ISO/IEC 11172。针对1/4的NTSC制和1/4的PAL制视像，压缩比为30:1时，相当于家用录像系统(VHS)的视像质量。针对采样频率为44.1kHz的声音数据，定义了3个层次的压缩比，压缩比约为6:1时，声音质量接近激光唱片上的声音质量。MPEG-1的主要应用：针对20世纪90年代初期CD-ROM 的数据传输能力只有1.4 Mb/s而开发的，MPEG-1的最高数据速率限定在1.5 Mb/s。MPEG-1标准主要用于在CD光盘上存储数字影视、在网络上传输数字影视以及存放MP3格式的数字音乐节目。

(2)**MPEG-1的系统模型：基本目标**：规范视像压缩和声音数据的编码标准。**MPEG-1Video**：支持子采样为4:2:0或4:1:1的视像。NTSC制: CIF格式，352像素/行×240行/帧×30帧/秒。PAL制：CIF格式，352像素/行×288行/帧×25帧/秒。经过压缩的视像输出速率为1.15 Mb/s；视像质量与VHS的视像质量相同。**MPEG-1Audio支持的声音：**采样频率最高为48 kHz，样本精度为16位。压缩后的数据速率分3个层次，最高速率分别为384, 256和192 kb/s。压缩后还原的声音质量接近于激光唱盘上的声音质量。**MPEG-1的总数据率控制在1.5 Mb/s左右**

**系统模型：**由编码系统和解码系统两大部分组成。编码系统由两个部分组成：(1) 视像编码和声音编码；(2) 系统层多路数据复合

2.**MPEG-2数字电视标准**

(1)**MPEG-2是什么:**1990年开始1994年完成的数字电视编码标准，(ISO/IEC 13818)。**基本编码算法与MPEG-1相同，增加了许多功能**：支持高分辨率的视像、大范围的数据速率、多声道的环绕声、多种视像分辨率、位速率可变(scalability)、隔行扫描等特性。**目标**：数据位速率最低4 Mb/s，最高100 Mb/s。**典型应用**：DVD影视和广播级质量的数字电视，包括美国的ATSC DTV、欧洲的DVB以及日本的ISDB，因特网电视。

(2)**MPEG-2的系统模型：**主要由编码系统和解码系统两部分组成。编码系统，由两个部分组成：(1) 视像编码和声音编码；(2)数据打包和多路数据复合。MPEG-2 Video规范支持的典型视像格式：NTSC制：分辨率为720像素/行×480行/帧×30帧/秒；PAL制：分辨率为720像素/行×576行/帧×25帧/秒。MPEG-2 Audio规范支持的声音：MPEG-1声音规范；高质量的环绕声，如5.1声道。把视像、声音和其他数据组合成适合存储或传输的基本数据流。**数据流有两种类型:**节目数据流：由基本数据流(PES)组合而成，用在出现错误相对比较少的环境，如DVD存储系统。传输数据流：由一个或多个PES组合而成，用在出现错误相对比较多的环境下，如电视广播、有损失或有噪声的传输系统。

**MPEG-2的类型与等级**:MPEG标准覆盖的应用范围很广，在不同环境下应用时，为避免增加具体应用系统的复杂性和浪费传输带宽，MPEG专家组为此引入了profile和level的概念，从标准中组合形成若干个profile，对每个profile指定不同的level。这套概念同样可以用在其他MPEG标准。**配置(profile)**：一个指定的语法子集，用于指定不同应用系统需要使用的算法，以及执行每种算法需要的工具。**等级(level)**：在特定profile下的一个限定取值的参数集，用于指定不同的计算复杂度和视像参数。**MPEG-2定义了5种类型的配置**，相应于5种类型的视像，分别是HP,SPP,SNRP,MP,SP。**视像的等级**：4个等级：按不同的视像分辨率和刷新频率定义，称为视像分辨率等级，(1) 低级(low level，LL)，公用中分辨率格式(CIF)(2) 基本级(main level，ML)，ITU-R BT601格式(3) 高级-1440(high-1440 level，H14L)，HDTV的一种格式(4) 高级(high level，HL)，HDTV的一种格式。

**为什么视像能压缩**?因为视像数据之所以能够被压缩，主要是视像数据中存在大量的冗余数据，包括时间冗余、空间冗余、结构冗余、视觉冗余、知识冗余和数据冗余。**MPEG视像标准利用的冗余：时间冗余、空间冗余、视觉冗余和数据冗余**。

**MPEG-1视像：**压缩数据的基本方法 在空间方向上，采用与JPEG类似的算法去掉空间冗余数据 在时间方向上，采用移动补偿(motion compensation)算法去掉时间冗余数据 MPEG专家组为此开发了两项重要技术：定义了视像数据的结构；定义了三种类型的图像。

**视像数据结构：**把视像片段看成一系列静态图像的序列。把视像序列分成许多像组(group of picture，GOP) 把像组中每一帧图像分成许多像片(slice)，每个像片由16行组成 把像片分成16行×16像素/行的宏块(macroblock) 把宏块分成若干个8行×8像素/行的图块(block)，见图2(a) 使用子采样格式为4:2:0时，一个宏块由4个亮度(Y)图块和两个色度图块(Cb和Cr)组成，

**三种类型的图像**：**(1)帧内图像I (intra- picture)**，简称I图像或I帧(I-picture / I-frame)。包含内容完整的图像，用于为其他帧图像的编码和解码作参考，称为关键帧。

**(2)预测图像P (predicted picture)**，简称P图像或P帧(P-picture / P-frame)。以在它之前出现的帧内图像I作参考的图像，对预测图像P进行编码就是对它们之间的差值进行编码。**(3)双向预测图像B (bidirectionally-predictive picture )**，也称双向插值图像B，简称B图像或B帧(B-picture/B-frame)。以在它之前和之后的帧图像(I和P)作参考的图像，对B进行编码就是对帧内图像I和预测图像P的差值分别进行编码。

**MPEG-2视像**

(1)MPEG-2视像是MPEG-1视像标准的扩展版本，全面继承了MPEG-1视像数据压缩算法，增添了许多新的语法结构和算法，用于支持：顺序扫描和隔行扫描；NTSC、PAL、SECAM和HDTV格式的视像；视像的实时传输。为适应各种不同的应用，MPEG-2视像标准定义了多种视像质量可变的编码方式。

(2)视像编码器和解码器，原理上与MPEG-1的编码和解码结构基本相同。“ME(移动估算器)”用于计算移动矢量，找出最佳匹配宏块。“内置解码器”用于产生预测图像，它的输入包括移动矢量、量化DCT系数和用于控制数据速率的量化参数控制信号。输入视像和预测图像通过“(加法器)”产生预测误差，经过“DCT(余弦变换)”和“Q(量化)”之后送给“VLE(可变长度编码器)”，移动矢量也送到“VLE”，它们在“VLE”经过编码和复合之后送到传输媒体或存储媒体。“量化参数控制”信号可改变视像质量和数据速率。

**MPEG-4视像**

（1）**MPEG-4 Visual是什么**：1999年发布的可视对象编码标准，第3版修改于2009年， ISO/IEC 14496-2 Part 2 **开发初衷**：针对不同应用，提供数据率小于64 kbps、64~384 kbps和0.384~4 Mbps的视像压缩编码，用于自然对象、合成对象以及合成-自然对象混合编码(SNHC)。**自然对象编码**：形状编码(shape coding)、纹理编码(texture coding)、移动编码(motion coding)和精灵编码(sprite coding)。**合成对象编码**：图形、人脸面部活动，身体动作等

**第十三章**H.264/AVC与H.265/HEVC

**H.264/AVC**:是H.264/ MPEG-4 AVC(MPEG-4 Part 10, Advanced Video Coding)的简写。是一种数字视像压缩标准。支持各种位速率和分辨率的电视图像，覆盖应用范围广泛,如电视广播、HD DVD和蓝光盘(Blu-ray)电视发行、网络电视、移动应用程序(APP)和多媒体电话系统。

**H.264/AVC提高编码效率的主要改进技术**:**[1]帧间预测**：采用可变图块的帧间预测和移动补偿，预测图块大小不局限于16×16像素，可小到4×4像素，提高移动矢量的预测精度。从过去编码后重构的相邻帧样本预测当前编码帧样本的过程。编码以块为基础。预测误差：用实际的样本值与预测值相减得到。对预测误差进行变换和编码，以消除时间方向上的冗余性**[2]帧内预测**：预测图块可以是16×16宏块，4×4图块，定义了多种预测方式，目的是找到匹配最佳的预测图块。在同一像片中从过去编码后重构的相邻图块对当前图块进行预测**[3]整数变换，**DCT演变而来，提高运算速度**[4]熵编码**：CAVLC/前后文自适应可变长度编码和CABAC /前后文自适应二元算术编码，效率比可变长度编码(VLC)高.**[5]多参考帧和消除块状失真的滤波技术**

**H.265/HEVC：**H.265/HEVC是在H.264/AVC基础上改进的视像编码标准，其结构和原理基本相同。主要目标：提高压缩比，改进并行处理方法，支持超高分辨率电视。需要的信号处理能力比H.264/AVC要高得多，但对解压缩需要的计算量的影响较小。

**第十四章MPEG音频编码标准**

声音的响度=声音的强弱。**人耳的听觉范围：**听阈：当声音弱到人耳朵刚可听见时的声音强度；痛域：声音强到使人耳感到疼痛时的声音强度；听觉范围：位于听阈和痛域之间。**对音调的感知**：客观上用频率表示声音的音高，单位为Hz。主观感觉的音高单位为美(Mel)。主观音高与客观音高的关系为Mel=1000\*log2(1+f)

**频率掩蔽效应**:一种频率的声音阻碍听觉系统感受另一种频率的声音的现象。前者称为掩蔽声音(masking tone)，后者称为被掩蔽声音(masked tone)，掩蔽可分成频域掩蔽和时域掩蔽。**时间掩蔽效应**：时间相邻的声音之间的掩蔽现象。一个强掩蔽音出现前、同时存在或消失后的掩蔽。同时掩蔽(simultaneous masking)：信号和掩蔽音同时产生的现象；滞后掩蔽(post-masking)：信号出现在掩蔽音消失后出现的现象，可以持续50～200 ms ；超前掩蔽(pre-masking)：信号出现在掩蔽音出现之前产生的现象。超前掩蔽很短，通常为2～20 ms。产生时域掩蔽的主要原因：人的大脑处理信息需要花费一定的时间。

**MPEG-1 Audio是什么?**

高保真的声音数据压缩国际标准(ISO/IEC 11172-3)。三个独立的压缩层次，一层比一层复杂。第1层—MP1(MPEG Audio Layer 1)：典型压缩比为4:1，数据率为384 kbps，应用包括小型数字盒式磁带(DCC)。第2层—MP2(MPEG Audio Layer 2)：典型压缩比为(6:1)～(8:1)，数据率为256～192 kbps，主要用于专业应用，如数字声音广播、数字音乐、VCD等。第3层—MP3(MPEG Audio Layer 3)：典型压缩比为(10:1)～(12:1)，数据率为128～112 kbps，在网上和网下都得到广泛应用，而且开发了许多MP3编码和播放设备。支持单声道、立体声和联合立体声

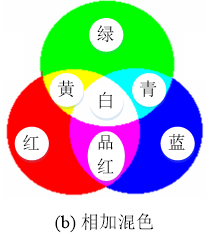
**MPEG-2 Audio简介**：MPEG-2 Audio和MPEG-1 Audio(ISO/IEC 1117-3)标准使用相同种类的编解码，3个编码层(第1, 2和第3层)的编码结构也相同。在MPEG-1声音基础上，做了如下扩充:①增加了16 kHz, 22.05 kHz和24 kHz采样速率②扩展了编码器的输出速率范围，由32～384 kbps扩展到8～640 kbps③增加了声道数，支持5.1声道和7.1声道的环绕声，支持Linear PCM(线性PCM)和Dolby AC-3(Audio Code Number 3)编码。**MPEG-AAC**:AAC= Advanced Audio Coding: 高级声音编码。是一种声音感知编码标准。使用掩蔽特性压缩声音数据，把量化噪声分散到各个子带，用全局信号来掩蔽噪声。采样频率：可从8 kHz到96 kHz，编码器的输入可来自单声道、立体声或多声道音源的声音。声道数：可支持48个，16个低频音效加强通道(LFE)，16个配音声道。声音质量：在11:1时，很难区分原始声音和重构声音。每个声道的数据率为(44.1×16 )/11=64 kbps，5个声道=320 kbps。压缩性能：在声音质量相同的前提下。与MPEG-1/-2 Audio 的第2层相比，AAC的压缩率可提高1倍。与MPEG-1/-2 Audio 的第3层相比，AAC的数据率是它的70％。

**MPEG-4 Audio简介**：编码对象：自然声音和合成声音。声音对象，目前已定义45种，如MP3、CELP。声音频率范围：20 ~ 20 000 Hz。技术：波形编码、参数编码、混合编码、感知编码。







**第一章 多媒体技术概要**

**超文本**：包含指向其他文档或文档元素的指针的电子文档

**超媒体**：包含指向文图声像文档或文档元素的指针的电子文档

**SGML(标准通用标记语言)**:精华是把文档内容与样式分开处理。主要特点：可支持无数的文档结构类型，如布告、技术手册、章节目录、设计规范、各种报告、信函和备忘录等可创建与特定软硬件无关的文档。很容易与使用不同计算机系统的用户交换文档

**多媒体层次结构**：基础层、系统层、服务层和应用层

**两种类型的压缩**

无损：压缩后的重构数据与原始数据相同

有损：压缩后的重构数据与原始数据不完全相同

**三种类型的编码**

熵编码：不考虑数据源特性的无损数据压缩技术

源编码：考虑数据源特性的数据压缩技术

混合编码：组合源编码和熵编码的数据压缩技术

**因特网**：通过网络设备把世界各地使用TCP/IP协议的计算机

相互连接在一起的计算机网络

**万维网**：在因特网上运行的全球性分布式多媒体信息系统。

万维网取得巨大成功的**三个关键标准**:URL,HTTP,HTML。

**制定多媒体标准的国际组织**：ITU(国际电信联盟) ,ISO/IEC(国际标准化

组织/国际电工技术委员会) -->制定了MPEG标准，IEEE (电气和电子工程师学会)  
制定了著名的局域网标准及其他标准,ETSI (欧洲电信标准学会)，ISOC(因特网协会)

负责协调因特网标准开发，W3C(万维网协会)

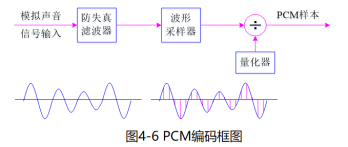
**DCIKW结构：**数据，内容，信息，知识，智慧。（->这个方向）

**标记语言**：标记内容的语言，是组织和表示数据的一套规则。

**两类标记：**程序标记(procedural markup)：

描述文档显示的样式(style)，如字体的大小、黑体、斜体和颜色等；描述标记

(descriptive markup)：描述文档中的文句，如篇、章、节、段落、列表或表格

**第二章 字符编码与字体**

**字符**：计算机使用的字母、数字和符号

**字符编码**：用代码表示字符的方法

**字体**：一整套字符的书写形式

**汉字编码三个基本问题**：汉字的编码，汉字的输入(输入法)，

汉字的显示。

**汉字字符编码方法**：码位(分配字符编号或称存放位置)

(字符)代码(分配一个数字)特点：汉字数量大。

**GB 2312标准是我国第一个汉字编码标准(1980)**

收录了7445个 = 6763个汉字 + 682个全角字符

使用双字节编码。

**区位码**：7445个字符组成94×94方阵，每一行称“区”，

编号0l～94；每一列称为“位”，编号0l～94

编码空间：94×94 = 8836个字符

**国标码**：区位码表示字符所在位置的编号，对这些编号的

编码就是国标码。

**编码方法**：先将十进制表示的区码和位码转换为十六进制表

示的区码和位码，再将这个代码的高字节(第一个字节)和低

字节(第二个字节)分别加上20H(10 0000B)，就得到国标码。

**机内码**：将国标码的高字节和低字节的最高位(b7)都变成1，

或者说每个字节都加0x80

**Unicode**：编码空间有17个平面，用5bit表示平面，每个

平面包含256行\*256码位/行=65536个码位，用十六进制

表示。

**UTF-8**用1~4个字节表示，**UTF-16**用1-2个字来表示，对

于在Unicode的基本多文种平面上定义的字符用一个字表

示，对于增补平面上的字符用两个字表示。

**两种显示字符**：点阵字符，矢量字符。

数字字体大致可分为**位图字体、轮廓字体和笔画字体。**

**位图字体：**由位图字符构成的字符集，每个字符用许多点(dot)或称像素(pixel)表示，1个字 = 1幅图。

**轮廓字体**：矢量字体，由轮廓字符构成的字符集。空心字体(hollow font)是轮廓字体中的一种

**笔画字体**：由笔画字符构成的字符集。字符使用各种规定的笔画和附加信息来定义，每种笔画的粗细、大小和外形用贝塞尔曲线来描述

**衬线字体：**在字符的笔画端处带短线或装饰线的字符构成的字符集；无衬线字体是字符的笔画端没有短线或装饰线的字符构成的字符集。中文的宋体字体是有衬线字符，黑体字符是无衬线字符。

**Type 1**使用三次贝塞尔曲线，**TrueType**使用二次贝塞尔曲线，**PostScript**字体使用三次贝塞尔曲线，**OpenType**字体包含65 536个字符，用Unicode编码。

**第三章 数据无损压缩**

**数据可被压缩的依据**：数据本身存在冗余,听觉系统的敏感度有限,视觉系统的敏感度有限

**三种多媒体数据类型**：文字 (text)数据——无损压缩，根据数据本身的冗余。声音(audio)数据——有损压缩，根据数据本身的冗余，根据人的听觉系统特性。

图像/视像数据——有损压缩，根据数据本身的冗余，根据人的视觉系统特性。**数据无损压缩的理论：信息论**

**数据无损压缩的方法**：霍夫曼编码，算术编码，行程长度编码，词典编码。**冗余的分类和概念**：**人为冗余**——在数据存储和传输中，为了检测和恢复在数据存储或数据传输过程中出现的错误，在数据存储或传输之前把额外的数据添加到数据中，这个额外的数据就是冗余数据，**视听冗余**——由于人的视觉系统和听觉系统的局限性，有些数据确实是多余的，删除后不会丢失实质性的信息或含义，**数据冗余**——不考虑数据来源时，单纯数据集中也可能存在多余的数据，去掉这些多余数据并不会丢失任何信息。**决策量**：在有限数目的互斥事件集合中，决策量为**H0=log(n)**,n是事件数。单位由对数的底数决定——log2->Sh,loge->Nat,log10->Hart

**信息量**：**I(x)=log2(1/p(x)]=-log2(p(x))**，p(x)为事件发生的概率。**熵**：在有限的互斥和联合穷举事件的集合中，熵为事件的信息量的平均值，也称事件的平均信息量。

计算公式为：。**冗余量**：数据的冗余量计算公式为决策量超过熵的量，公式为R = H0-H

**统计编码**：是一种给已知统计信息的符号分配代码的数据无损压缩方法。**编码方法有**：香农-范诺编码，霍夫曼编码，算术编码。**编码特性**：香农-范诺编码和霍夫曼编码的原理相同，都是根据符号集中各个符号出现的频繁程度来编码，出现次数越多的符号，给它分配的代码位数越少算术编码使用0和1之间的实数的间隔长度代表概率大小，概率越大间隔越长，编码效率可接近于熵。

**香农-范诺编码**：熵的大小表示非冗余的不可压缩的信息量，他是一种从上到下的编码方式。**具体算法**：就是用香农公式计算H(X)，然后求压缩比。使用该编码方式来进行压缩，将所有数据按照出现频率降序排列，然后从上到下的进行递归的划分，要求使得两边大小之差最小。根据这种编码之后再求压缩比。

**霍夫曼编码**：是一种从下到上的熵编码方法。根据给定数据集中各元素所出现的频率来压缩数据的一种统计压缩编码方法。这些元素(如字母)出现的次数越多，其编码的位数就越少。广泛用在JPEG, MPEG, H.26X等编码标准中。**具体算法**：将所有元素按照出现次数降序排序，然后根据次数构建一个哈夫曼树，再在上面标上01，一般次数多的标1，由此得到各元素的编码值。**计算平均码长的公式**：，其中li为第i个元素的编码的长度，p(li)为其出现的频次。

**算术编码**：基本思想是用0和1之间的一个数值范围表示输入流中的一个字符，而不是给输入流中的每个字符分别指定一个码字，实质上是为整个输入字符流分配一个“码字”，因此它的编码效率可接近于熵。**具体算法**：先初始化，将[0,1)根据其中元素的频率划分为几段，然后再遍历待压缩的内容，根据是哪个元素选择哪一段，按照上述方法再划分选中的那一段区间。最终得到的一个区间，选择其中的一个数字作为编码值。解码就是根据这个编码值，看他属 于哪个区间，属于哪个原来就是这个值，然后再进行区间划分，再看属于哪个区间。

**RLE编码**：一种无损压缩数据编码技术，它利用重复的数据单元有相同的数值这一特点对数据进行压缩。对相同的数值只编码一次，同时计算出相同值重复出现的次数。在JPEG，MPEG，H.261和H.263等压缩方法中，RLE用来对图像数据变换和量化后的系数进行编码。

**词典编码**：文本中的词用它在词典中表示位置的号码代替的一种无损数据压缩方法。采用静态词典编码技术时，编码器需要事先构造词典，解码器要事先知道词典。采用动态辞典编码技术时, 编码器将从被压缩的文本中自动导出词典，解码器解码时边解码边构造解码词典。**分类**：**第一类编码算法**——用已经出现过的字符串替代重复的部分，编码器的输出仅仅是指向早期出现过的字符串的“指针” 。**第二类编码算法**——从输入的数据中创建一个“短语词典“，编码器输出词典中的短语“索引号”，而不是短语。**具体算法有**：LZ77算法，LZSS算法，LZ78算法，LZW算法。**LZW算法原理**：lzw算法是围绕称为词典的转换表来完成的，其中用来存放称为前缀的字符序列，并且为每个表项分配一个码字或者叫序号。

**霍夫曼编码和算数编码的比较**：算术编码实现复杂度更高，压缩效率要比霍夫曼稍好些，专利保护多；霍夫曼编码实现复杂度低（开销更小），压缩效率略逊于算术编码，保护专利较少；非对称数字系统（ANS）编码（2014以后出现的新方法），两者兼顾，压缩效率高，复杂度低。

**第四章 数字语音编码**

**语音**：频率为300~3400 Hz的信号，**声音**：频率为20~ 20 000Hz的信号，**语音是声音，但声音不一定是语音**。**常见语音编码**：波形编码，参数编码(LPC为主)，混合编码(以CELP为主)。**声音名称和频率范围：**高保真声音: 10 ~ 20 000 Hz，声音: 20~ 20 000 Hz，语音: 300~3400 Hz，亚音: < 20 Hz，超声: > 20 000 Hz，人的发音器官发出的声音频率大约在80～3400 Hz，男音：300～3000 Hz，女音：300～3400 Hz，语音：300～3400 Hz。**模拟信号**：时间和幅度上都是连续的信号。**数字信号**：时间和幅度都用离散值表示的信号。在特定时刻对模拟信号进行测量叫做**采样**(sampling)，由这些特定时刻采样得到的信号称为离散时间信号。采样得到的幅值是无穷多个实数值中的一个。如果把信号幅度取值的数目加以限定，这种由有限数目的数值组成的信号称为**离散幅度信号。采样速率/采样频率：**由信号本身包含的最高频率决定，信号的频率越高，需要的采样速率就越高。**奈奎斯特理论：**采样速率不应低于声音信号最高频率的两倍。

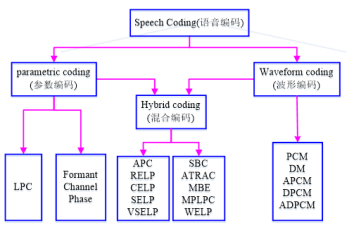
**奈奎斯特速率**：fs>=2\*f。其中f是声音信号包含的最高频率，fs表示采样速率。当满足这个不等式的时候，fs就称为奈奎斯特速率。而速率的一半（fs/2)就被称为奈奎斯特频率。**使用奈奎斯特理论采样**，能把以数字表达的声音还原成原来的声音，称为**无损数字化。**

**采样精度**：度量声音波形幅度的精确程度，通常用存储每个声音样本的位数(n)来表示——e.g 例如，每个声音样本用16位(2字节)表示，测得的声音样本值是在[0～65535]范围里的数，采样精度是1/65536。**精度是度量模拟信号的最小单位**，称为**量化阶**（也就是采样精度）。**位数越多，声音质量越高，所需存储空间也越多。**

**信噪比**：采样精度的另一种表示方法是信号噪声比，简称为信噪比。公式为：SNR = 20lg(V\_signal/V\_noise)，假设采样精度为n位，则SNR = 20lg(V\_signal\*2^n/V\_signal)

=20lg(2^n)=20nlg2。根据声音频带，**声音质量**分成**5个等级**，由低到高：电话，调幅(AM)广播声音，调频(FM)广播声音，激光唱盘声音，数字录音带声音。

**声音质量的MOS评分标准**：声音信号的带宽，客观质量度量：主要用信噪比(SNR)，主观质量度量：人的感觉更具决定意义——5分制，分为优良中差劣。

**波形编码**：用数字形式精确地表示模拟信号波形的编码方法，不考虑语音产生和感知特性。编码器输出的数据速率为9.6~64 kbps或更高(取决于采样速率和量化精度)。是语音编码质量最好的编码方法。

**参数编码**：利用发音器官生成语音信号的模型，对从语音信号中抽出的语音特征参数(如发音模型、有声/无声、音量大小、音调)进行编码的方法。解码器根据模型参数重构语音信号。编码器输出的数据速率约为2~4.8 kbps。

**混合编码**：综合使用波形编码和声源编码技术，组合波形特性和语音特征参数的编码方法。既有波形编码语音质量高的优点，又有参数编码数据速率低的优点。主要用在语音质量要求高的移动通信系统。混合编码器输出的数据速率为4.0～16 kbps。

**PCM编码**：脉冲编码调制PCM。概念最简单、理论最完善的编码系统，最早研制成功、使用最为广泛的编码系统，数据量最大的编码系统，1972年ITU将它作为G.711声音(audio)编码标准。

**防失真滤波器**：低通滤波器，滤除声音频带外的信号。

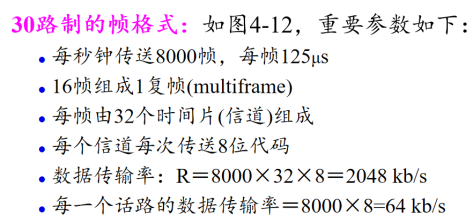
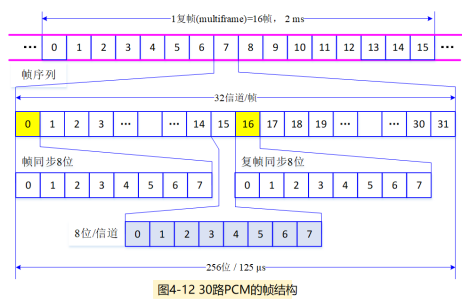
**均匀量化（线性量化）**：采用相等的量化间隔对采样信号进行量化。量化误差或量化噪声：样本值Y和原始值X之差 E=Y-X。  
**非均匀量化**：对幅度大的信号采用大的量化间隔，幅度小的信号采用小的量化间隔。可在满足精度要求的情况下用较少的位数来表示。有两种算法：µ律压扩和A律压扩。这两种算法与人的听觉感知特性一致。

**µ律压扩**：按照如下公式确定量化输入输出的关系。F(x)=sgn(x)\*ln(1+μ\*|x|)/ln(1+μ)，其中x为输入信号幅度，规格化到[-1,1]，函数sgn(x)的值为 1，符号同x的正负号;μ为确定压缩量的参数，它反映最大量化间隔和最小量化间隔之比，取 100≤μ≤500。由于μ律压扩的输入和输出关系是对数关系，所以这种编码又称为对数PCM。具体计算时，用μ=255，把对数曲线变成8条折线以简化计算过程。

**A律压扩**：按照如下公式确定量化输入输出的关系。当0<=|x|<=1/A, F(x)=sgn(x)\*A\*|x|/(1+lnA),当1/A<|x|<=1,F(x)=sgn(x)\*(1+ln(A\*|x|)/(1+lnA)。式中，x为输入信号幅度，规格化成-1≤x≤1;sgn(x)为x的极性;A为确定压缩量的参数，反映最大量化间隔和最小量化间隔之比。A律压扩的前一部分是线性的，其余部分与u律压扩类似。具体计算的时候，A=87.56，为简化计算，同样可以把对数曲线部分变成折线。**PCM编码用在时分多路复用的语音通信中**

**时分多路复用**(TDM)：在同一条通信线路上使用不同时段“同时”传送多个独立信号，核心思想：将时间分成等间隔的时段，为每个用户指定一个时间间隔，使许多用户同时使用一条传输线路。**语音信号的采样速率f＝8000 Hz/s，采样周期＝125 s，称为1帧(frame)**。在1帧时间里，可容纳的话路数有两种规格——24路制，30路制。



**数字通信线路的数据传输率：** 群(group)：反映PCM信号复用的复杂程度，也称为数字网络的等级。一次群(基群)：30路或24路，二次群：120路或96路。



**增量调制(DM)**:一种预测编码技术。对实际的采样信号与预测的采样信号之差的极性进行编码，差值为正用1表示，差值为负用0表示。

**自适应增量调制(ADM)**：根据输入信号斜率变化自动调整量化阶Δ的大小

**APCM(自适应脉冲编码调制)**：根据输入信号幅度或音节来改变量化阶的大小。自适应：量化阶的大小每隔几个样本改变一次，音节自适应(量化阶的大小在较长时间周期里发生变化)。改变量化阶大小的两种方法：前向自适应(forward adaptation)：根据未量化的样本值的均方根估算输入信号的幅度，以此来确定量化阶大小，并对其幅度进行编码作为边信息(side information)；后向自适应(backward adaptation)：从量化器刚输出的过去样本中来提取量化阶信息，不需要传送边信息。

**预测编码**：预测编码(predictive coding)是对实际样本值与预测值之差进行编码的方法。在信道上传送的是相邻样本之间的差值，解码时使用差值和预测值进行重构，预测函数一般为x\_pre(n)=。其中an-k为预测系数。

**DPCM(差分脉冲编码调制)**：利用样本与样本之间存在冗余进行编码的数据压缩技术，基本思想：根据过去的样本估算下一个样本信号幅度，称为预测值，然后对实际信号值与预测值之差进行量化编码。

**ADPCM(自适应差分脉冲编码调制)**：综合APCM的自适应特性和DPCM的差分特性。核心想法：利用自适应改变量化阶大小，用小的量化阶(step-size)编码小的差值，用大的量化阶编码大的差值，使用过去的样本值估算当前输入样本的预测值，使实际样本值和预测值之间的差值总是最小

**音频存储量计算公式**：存储量 (Byte)=采样频率 (Hz)×量化精度 (bit)/8×声道数×时间 (s)

**第五章 彩色数字图像基础**

**图像数据压缩**：图像数据中有许多重复的数据，用无损数据压缩技术；人眼对图像细节和颜色的辨认有极限，用有损数据压缩技术。

可见光是波长在**380～780 nm**之间的电磁波。**描述颜色常用的两个互为同义术语**：**颜色模型**(color model)：用数值指定颜色的方法；**颜色空间**(color space)：用空间中点的集合描述颜色的方法。**RGB和CMYK是计算机系统广泛使用的颜色模型。相加混色法：**组合三种光波产生特定颜色的方法，也称RGB相加混色模型。**彩色图像**：看成是由许多点组成的。像素(pixel)：单个像点。像素值：颜色强度，常用R，G，B三个颜色分量表示。若三个颜色分量都用1位表示，每个分量强度只有1(100%)和0(0%)，可显示8种颜色之一。**VGA适配卡**：每种颜色用8位表示，可产生2^24=16 777 216种颜色。Windows系统用代码0～15表示，1～6：颜色较暗，用最大光强值的一半产生，9～15：用最大光强值产生。

**HSL模型(色调-饱和度-亮度模型)**：H定义颜色的波长，称为色调，S定义颜色的强度(intensity)，表示颜色深浅，称为饱和度，L定义掺入的白光量，称为亮度。重要性：比较容易为画家理解，若把S和L值设置为1，改变H时选择不同的纯颜色，减小饱和度S时，体现掺入白光的效果，降低亮度时，颜色就暗，相当于掺入黑色。**打印彩色图像用CMY相减混色模型：**无源物体: 不发光波的物体，颜色由该物体吸收或反射哪些光波决定，用CMY相减混色模型

用彩色墨水或颜料进行混合，绘制的图画是无源物体。**青色、品红和黄色，通常写成CMY**，称为CMY模型。用这种方法产生的颜色称为相减色，因为它减少了为视觉系统识别颜色所需要的反射光。

彩色打印和印刷彩色图片采用相减混色模型，由于彩色墨水和颜料的化学特性，用等量的三基色得到的黑色不是真正黑色，在印刷术中常加真正的黑色(black ink)。CMY写成CMYK。

**图像的三个基本属性**：分辨率，像素深度，真/伪彩色。

**图像分辨率**：屏幕分辨率和图像分辨率。**屏幕分辨率/显示分辨率**：衡量显示设备再现图像时所能达到的精细程度，常用水平和垂直方向所能显示的像素数目表示，写成“水平像素数×垂直像素数”。宽高比：传统的水平分辨率与垂直分辨率之比为4∶3；高清晰度电视和显示器的宽高比为16∶9。**图像分辨率**：是图像精细程度的度量方法。图像分辨率也称空间分辨率。在图像显示应用中，**图像分辨率有多种方法表示**：

物理尺寸：每毫米线数(或行数)

行列像素：像素/行×行/幅，如640像素/行×480行/幅

像素总数，如手机相机标的“1600万像素”

单位长度像素：如像素每英寸(pixels per inch，PPI)

线对数：以黑白相邻的两条线为一对，如“每毫米10线”表示黑线和白线相间的5对线。还有就是像素深度。

在图像数字化和打印应用中，用**每英寸像素点(dots per inch，DPI)**表示。e.g用300 DPI来扫描一幅8″×10″的彩色图像，得到一幅2400×3000个像素的图像。分辨率越高，像素越多。图像分辨率与屏幕分辨率是两个不同的概念。图像分辨率是构成一幅图像的像素数目，而屏幕分辨率是显示图像的区域大小。

**像素深度**：存储每个像素所用的位数。e.g在图像信号数字化时，记录每个图像样本信号的位数为8，10，12或16位。8位表示分辨率为1/256，10位表示分辨率为1/1024。像素深度也被认为是图像分辨率的度量方法。像素深度决定彩色图像的每个像素可能显示的颜色数，或灰度图像的每个像素可能显示的灰度级数。像素深度或图像深度可以很深，但各种VGA显示的颜色深度却受到限制。【例】标准VGA支持4位16种颜色的彩色图像，多媒体应用中通常推荐用8位256种颜色。

**α通道**：表示彩色图像像素时，除R，G，B分量用固定位数外，往往还增加1位或几位作为属性(attribute)位。例如，RGB 5∶5∶5表示一个像素用16位表示，其中R，G，B各占5位，剩下最高1位(b15)作为属性位，用来指定该像素具有的性质，称为透明(transparency)位，记为T。T的含义：假如显示屏上已经有一幅图存在，如果要把另一幅图重叠在它上面，可用T位控制原图是否能看得见。【例】T=1，原图完全看不见；T=0，原图能完全看见。属性位T称为1位α 通道(alpha channel)，像素深度为16位，图像深度为15位。**通道可看作预乘数通道**。【例】一个像素(A，R，G，B)的4个分量均用规一化值，当像素值为(0.5，1，0，0)时，使用预乘数0.5与R，G，B相乘，结果为(0.5, 0.5, 0, 0)，表示显示的红色强度为0.5。用两幅图像A和B混合成一幅新图像，新像素为，New pixel =(alpha)(pixel A color) +(alpha)(pixel B color)

**α通道的实际作用**：在彩色图像上叠加文字，不让文字覆盖图像，可用α通道；把该像素显示的颜色称为混合色(key color)；把电视图像和计算机生成的图像混合称为视图混合(video keying)。

**真彩色**：每个像素的颜色值用R，G，B表示的颜色。真彩色的颜色数目：2^24 ＝16777216种。通常用24位表示，称为24位颜色或全彩色(full color)。

**伪彩色**：每个像素的颜色不是由每个基色分量数值直接决定的颜色，而是把像素值当作彩色查找表(color look-up table，CLUT)的表项入口地址，去查找显示图像时使用的R，G，B值，用查找出的R，G，B值产生的彩色称为伪彩色。彩色查找表(CLUT)：事先做好的表，表项入口地址称为索引号。

在计算机中常见两种类型的图：矢量图(vector graphics)和位图(bitmap或bitmapped image)。

**矢量图**：用一系列计算机指令描绘的图，如点、线、面、曲线、圆、矩形以及它们的组合。绘制和显示这种图的软件通常称为绘图程序，存放这种图的存储格式称为矢量图格式，存储的数据主要是绘制图形的数学描述。优点：容易移动图像、缩小或放大、旋转、拷贝、改变属性(如线条变宽变细、颜色)，相同或类似图可当作图的构造块，矢量图的文件小。但难以用数学方法描述真实世界的彩照，需用位图表示。

**位图**：用像素值阵列表示的图。不论用什么方法生成的图，凡是用像素值阵列表示的图都称为位图。只能对图中的像素进行操作。表示矢量图的位图称为图形图像(graphical image)或矢量图像(vector based image)，也称光栅图(raster graphics)。画图或编辑位图的软件称为画图程序(paint programs)，存放位图的格式称为位图格式，存储的图像是像素的数值。位图的获取通常用扫描仪、数码相机、摄像机、录像机、视像光盘和相关的数字化设备。位图文件占据的存储空间大，文件大小由图像分辨率和像素深度决定。

**灰度图**：只有明暗不同的像素组成的图像。单色图像(bit image)：只有黑白两种颜色，每个像素的像素值用一位存储，其值是0或1。灰度图：如每个像素的像素值(0～255)用8位表示，灰度级数等于256。

**彩色图像**：按颜色数目划分，如256色图像和真彩色图像。

**γ的概念**：大多数光电转换特性都是非线性的，但有一个能够反映各自特性的幂函数，一般形式为y=x^n => 输出=（输入）^γ。γ(gamma)：衡量非线性部件的转换特性，称为幂-律(power-law)转换特性。输入和输出缩放到0～1之间，0为黑电平，1为颜色分量的最高电平。对特定部件，可度量它的输入与输出之间的函数关系，从而找出γ值，如投影幻灯片的γ值为1.5左右。如果整个图像系统的γ＝1，可再现真实的原始场景。

**γ校正**：所有显示设备都有幂-律转换特性，如果生产厂家不加说明，它的γ值大约等于2.5。使整个系统的γ值接近于使用要求，需要可校正γ值的非线性部件，用来校正显示设备的非线性。在所有广播电视系统中，γ校正在摄像机中完成。

**JPEG：**Joint Photographic Experts Group，静态数字图像数据压缩标准。JPEG专家组开发了两种基本的压缩算法：采用以离散余弦变换(Discrete Cosine Transform，DCT)为基础的有损压缩算法。压缩比为25:1时，非图像专家难以找出压缩前后图像之间的的差别；采用预测技术为基础的无损压缩算法。JPEG专家制定了采用小波变换的JPEG 2000标准。**JPEG算法概要**：利用视觉系统特性，使用变换、量化和熵编码相结合，去掉或减少视觉冗余和数据冗余。压缩算法大致分成三个步骤：1.用正向离散余弦变换(forward discrete cosine transform，FDCT)把空间域表示的图变换成频率域表示的图2.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编码器对量化系数进行编码。**颜色空间变换(RGB到YUV)和(YUV到RGB)不包含在JPEG算法**，处理单独的彩色分量图像。JPEG算法的计算步骤：(1) 正向离散余弦变换(FDCT)(2) 量化(quantization)(3) Z字形编码(zigzag scan)(4) 使用DPCM对直流系数(DC)进行编码(5) 用行程长度编码(RLE)对交流系数(AC)进行编码(6) 熵编码(entropy coding)。

**SVG(Scalable Vector Graphics)**:可缩放矢量图形。用XML格式描述二维(2D)矢量图形的语言。用SVG语言创建的图称为SVG图形，创建的文件称为 SVG文件，其扩展名为.svg，使用gzip压缩的文件扩展名为.svgz。SVG图形用XML格式的文字描述，生成的文件小，容易被检索，放大缩小都不会损失图的质量，现在的浏览器也都支持SVG文件，因此已成为像GIF和JPEG那样的标准，在因特网和移动设备上逐渐流行。SVG可支持光栅图像。两种方法:把图像文件嵌入(embed)到SVG文档，存储为单个SVG文件；把图像链接(link)到SVG文档，存储的SVG文件只含图像的绝对地址或相对地址，不包含图像本身

**文件格式**：存储文本、图形或图像数据的数据结构。**图像文件格式**有：BMP——位图文件格式，扩展名bmp,GIF——采用LZW算法压缩图像数据，允许用户为图像设置透明的背景，可在一个文件中存放多幅彩色图形/图像，可像幻灯片那样显示或像动画那样演示。,JPEG——颜色空间：使用1982年推荐的电视图像信号数字化标准CCIR 601，更名为ITU-R BT.601,PNG——位图文件(bitmap file)存储格式。存储灰度图像时，灰度图像的深度可多达16位。存储彩色图像时，彩色图像的深度可多达48位，还可存储多达16位的α通道数据。使用从LZ77派生的数据无损压缩算法。PNG文件格式保留了GIF文件格式的许多特性。

**图像存储量计算公式**：存储大小（字节）= 宽度（像素）x 高度（像素）x 色彩深度（位/像素）/ 8（转换为字节）



