

# 多媒体技术基础

## 提纲

- 数字图像压缩编码
- JPEG图像压缩（掌握）
- JPEG2000图像压缩（了解）
- 基于神经网络的图像压缩（了解）

授课教师：钱学明

西安交通大学

qianxm@mail.xjtu.edu.cn

SMILES LAB, XJTU

<http://smiles-xjtu.com/>

## 1. 图像颜色空间转换

RGB空间RGB分量的相关性强；存在大量的冗余；YUV空间中YUV相关性弱，冗余小



## 2. 图像压缩必要性及出发点

① 为何要进行数字图像压缩？

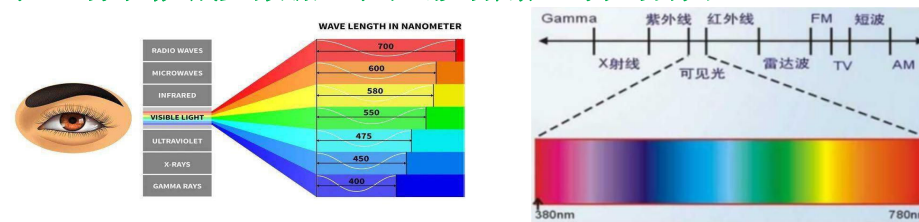
$512 \times 512 \times 3 \times 8 = 876.432 \text{ KBytes} = 6,291,456 \text{ bits}$

80年代计算机系统：内存128K~1M, CPU 5MHz~100MHz

② 从哪些方面进行压缩？

1) 人眼的视觉特性及视觉冗余

视觉冗余是人眼无法察觉到的图像变化，对于这种无法引起视觉敏感的信息可以适当舍弃，系统会对视觉敏感和不敏感的部分进行划分，对于那些无法引起视觉感知的变化则进行忽略，这样既能减少数据量，又影响用户的观看体验。



## 2. 图像压缩必要性及出发点

### ② 从哪些方面进行压缩?

#### 2) 图像内容冗余: 空间冗余 & 数据统计冗余

画面中空间邻近区域具有颜色、纹理相似性、边缘一致性等



示例图像1



示例图像



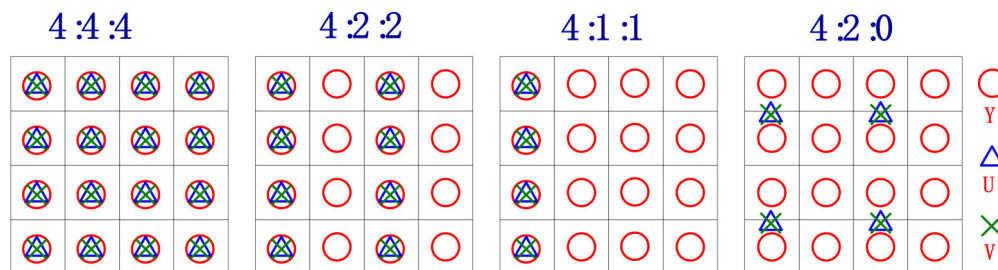
局部图像区域及像素值

## 3. 图像压缩标准

### ③ 有哪些标准进行压缩?

JPEG	基于DCT变换的压缩	(掌握)	1980s
JPEG2000	基于小波变换的压缩	(了解)	2000s
基于深度神经网络的压缩		(了解)	2020s

#### ➤ 色度分量下采样策略



## 3. 图像压缩的标准

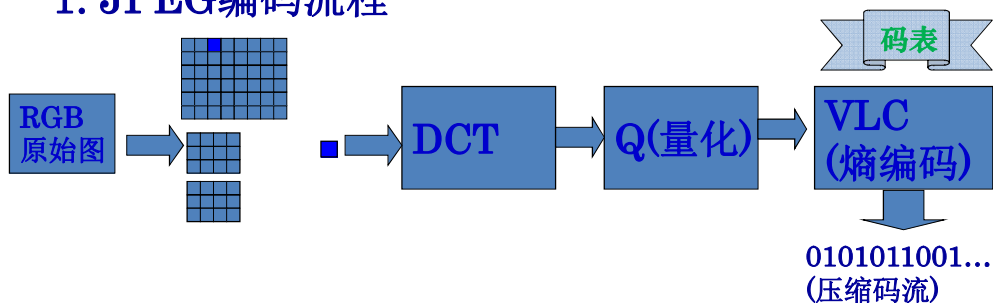
### ➤ JPEG压缩标准

一个H\*W\*3, RGB图像怎么压缩?

视觉冗余  
空间冗余  
统计冗余



### 1. JPEG编码流程



## 3. 图像压缩的标准

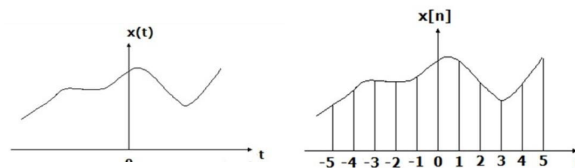
### ➤ JPEG压缩标准

#### 1维DFT变换

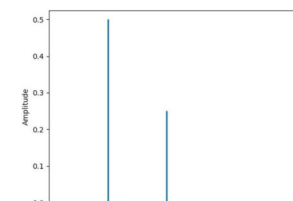
$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j(2\pi/N)kn}, \text{ 其中 } (0 \leq k \leq N-1)$$

原信号  $x(t)$  的采样信号  $x[n]$  可以用  $X[k]$  表示为:

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{j(2\pi/N)kn}, \text{ 其中 } (0 \leq n \leq N-1).$$



DFT变换的作用: 获得信号稀疏的描述



### 3.图像压缩的标准

#### ➤JPEG压缩标准

##### ✓ 图像分块DCT变换

$$F(u) = c(u) \sum_{i=0}^{N-1} f(i) \cos \left[ \frac{(i+0.5)\pi}{N} u \right]$$

$$c(u) = \begin{cases} \sqrt{\frac{1}{N}}, & u = 0 \\ \sqrt{\frac{2}{N}}, & u \neq 0 \end{cases}$$



##### ✓ 2维DCT变换

$$F(u,v) = c(u)c(v) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i,j) \cos \left[ \frac{(i+0.5)\pi}{N} u \right] \cos \left[ \frac{(j+0.5)\pi}{N} v \right]$$

$$c(u) = \begin{cases} \sqrt{\frac{1}{N}}, & u = 0 \\ \sqrt{\frac{2}{N}}, & u \neq 0 \end{cases}$$

##### ✓ 2D维DCT变换的矩阵形式

$$F = AfA^T$$

$$A(i,j) = c(i) \cos \left[ \frac{(j+0.5)\pi}{N} i \right]$$

### 3.图像压缩的标准

#### ● DCT量化矩阵

##### Y-分量

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

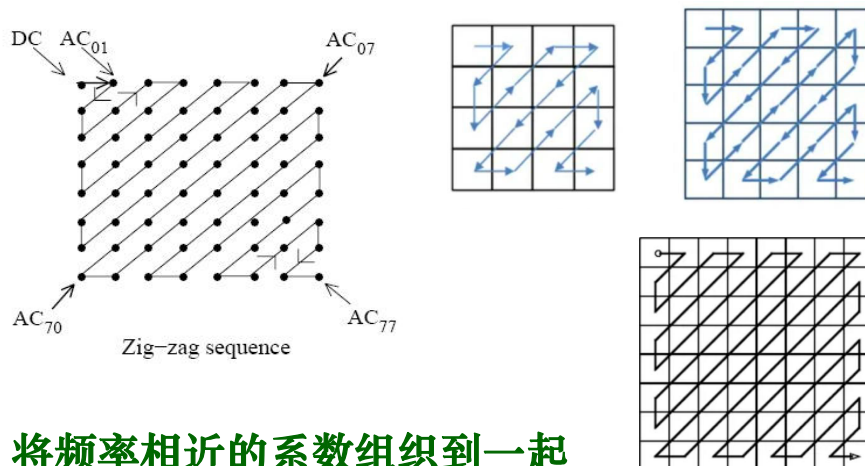
##### UV-分量

$$Q(u,v)_{UV} = \begin{bmatrix} 17 & 18 & 24 & 47 & 99 & 99 & 99 & 99 \\ 18 & 21 & 26 & 66 & 99 & 99 & 99 & 99 \\ 24 & 26 & 56 & 99 & 99 & 99 & 99 & 99 \\ 47 & 66 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \end{bmatrix}$$

根据量化矩阵可以确定量化参数  
充分体现人眼对色度分量不敏感的特点

### 3.图像压缩的标准

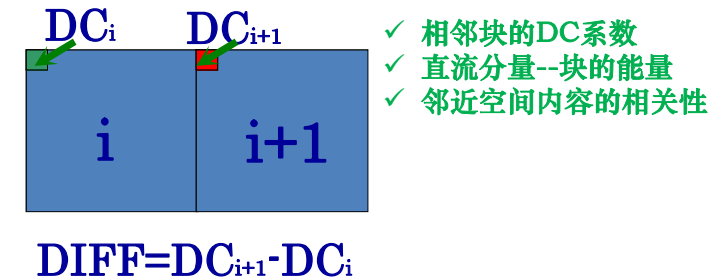
#### ● DCT系数重排序-ZigZag扫描



将频率相近的系数组织到一起

### 3.图像压缩的标准

#### ●DC系数编码DPCM



问题：为何仅做DC系数的预测？



### 3.图像压缩的标准

#### ● VLC熵编码:

无损编码，消除统计冗余

#### 例子: Huffman 编码原理

将符号出现频度(概率)大的像素值给一个比较短的编码，将出现频度小的像数值给一个比较长的编码。

Example: efabfcecafedffbefafbea

Stat. a: 4; b:3; c: 2; d:1; e:5; f:7.

编码结果?

11011000101000111000110011100000101001110110010011110

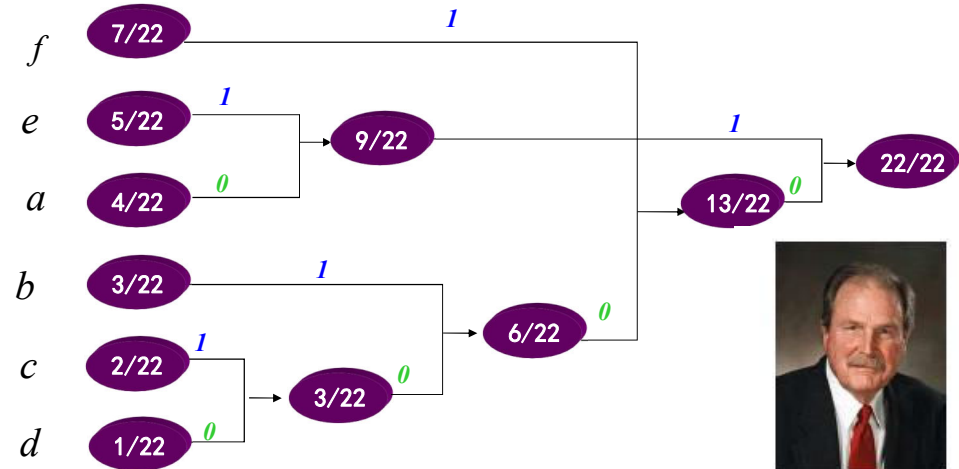
编码bits:  $4*2+3*3+2*4+1*4+5*2+7*2=56$

Huffman编码结果 f:01 e:11 a: 10 b:001 c:0001 d: 0000

讨论: VLC的好处有哪些? 在存储和传输中所导致的问题有哪些?

### 3.图像压缩的标准

熵编码原理--Huffman 编码原理示例

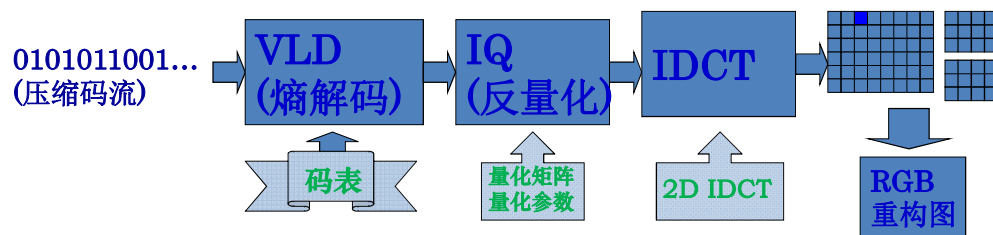


David Huffman

f=01 e=11 a=10 b=001 c=0001 d=0000

### 3.图像压缩的标准

#### JPEG解码流程



8\*8块  $F(u,v)$  的IDCT:

$$f(x,y) = \frac{1}{4} \left[ \sum_{x=0}^7 \sum_{y=0}^7 C(u)C(v) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} F(u,v) \right]$$

$$c(u) = \begin{cases} \sqrt{\frac{1}{N}}, & u = 0 \\ \sqrt{\frac{2}{N}}, & u \neq 0 \end{cases}$$

$$F = AfA^T$$

$$A^{-1} = A^T$$

$$f = A^{-1}F(A^T)^{-1} = A^TFA$$

### 3.图像压缩的标准

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

2维图像块

$$g(i,j) = \begin{bmatrix} 200 & 202 & 189 & 188 & 189 & 175 & 175 & 175 \\ 200 & 203 & 198 & 188 & 189 & 182 & 178 & 175 \\ 203 & 200 & 200 & 195 & 200 & 187 & 185 & 175 \\ 200 & 200 & 200 & 200 & 197 & 187 & 187 & 187 \\ 200 & 205 & 200 & 200 & 195 & 188 & 187 & 175 \\ 200 & 200 & 200 & 200 & 200 & 190 & 187 & 175 \\ 205 & 200 & 199 & 200 & 191 & 187 & 187 & 175 \\ 210 & 200 & 200 & 200 & 188 & 185 & 187 & 186 \end{bmatrix}$$

去直流分量

$$f(i,j) = \begin{bmatrix} 72 & 74 & 61 & 60 & 61 & 47 & 47 & 47 \\ 72 & 75 & 70 & 60 & 61 & 54 & 50 & 47 \\ 75 & 72 & 72 & 67 & 72 & 59 & 57 & 47 \\ 72 & 72 & 72 & 72 & 69 & 59 & 59 & 59 \\ 72 & 77 & 72 & 72 & 67 & 60 & 59 & 47 \\ 72 & 72 & 72 & 72 & 72 & 62 & 59 & 47 \\ 77 & 72 & 71 & 72 & 63 & 59 & 59 & 47 \\ 82 & 72 & 72 & 72 & 60 & 57 & 59 & 58 \end{bmatrix}$$

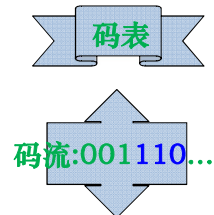
量化矩阵

2维DCT变换[方便显示取整]

$$F(u,v) = \begin{bmatrix} 515 & 65 & -12 & 4 & 1 & 2 & -8 & 5 \\ -16 & 3 & 2 & 0 & 0 & -11 & -2 & 3 \\ -12 & 6 & 11 & -1 & 3 & 0 & 1 & -2 \\ -8 & 3 & -4 & 2 & -2 & -3 & -5 & -2 \\ 0 & -2 & 7 & -5 & 4 & 0 & -1 & -4 \\ 0 & -3 & -1 & 0 & 4 & 1 & -1 & 0 \\ 3 & -2 & -3 & 3 & 3 & -1 & -1 & 3 \\ -2 & 5 & -2 & 4 & -2 & 2 & -3 & 0 \end{bmatrix}$$

量化后的系数[待编码]

$$\hat{F}(u,v) = \begin{bmatrix} 32 & 6 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$



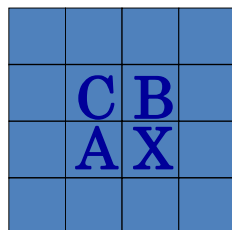
### 3.图像压缩的标准



### 3.图像压缩的标准

#### 有损JPEG&无损JPEG编码

- 有损JPEG编码 流程
    - 基于分块DCT变换
  - 无损JPEG编码方法
    - 采用邻域预测 + Huffman编码
- 当前像素X用ABC像素值进行预测



- 预测:  $P = A + B - C$
- 残差:  $X - P$  Huffman编码

比较:  
一般无损JPEG编码的压缩率只有2  
而有损的JPEG编码的压缩比很高

### 3.图像压缩的标准



### 4.本章小结

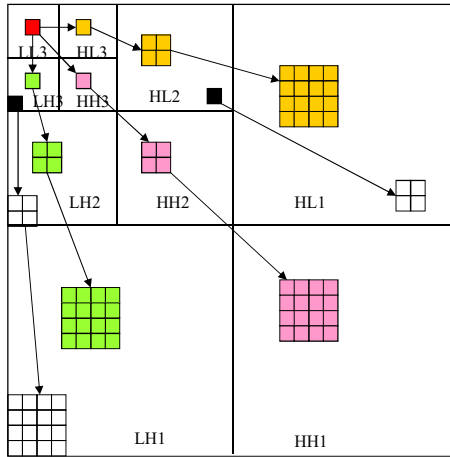
图像成像原理简要介绍  
图像视频中的冗余信息  
有损JPEG&无损JPEG编码

参考材料:  
<https://www.keyence.com.cn/ss/products/vision/visionbasics/>

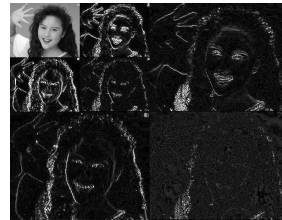
## 5.JPEG2K图像压缩标准

### ➤JPEG2K图像压缩

基于分块DCT变换的方法存在的问题及局限性

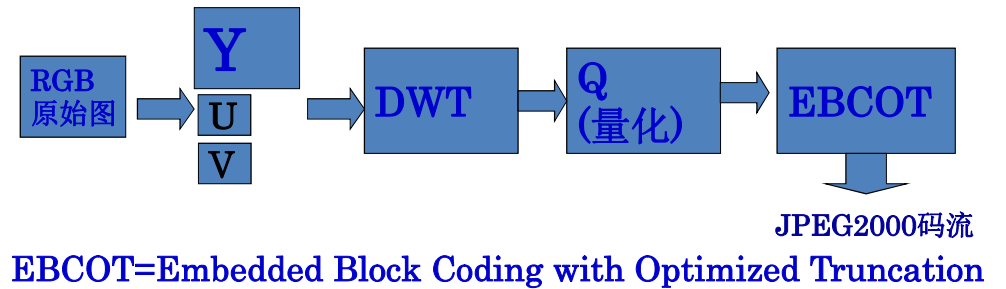


能量分布统计:  
CA:234.86  
CH:3.73  
CV:3.52  
CD:4.52



## 5.JPEG2K图像压缩标准

### ➤ JPEG2000编码流程



## 6.CNN图像压缩方法

### ➤卷积神经网络 (CNN)

基于卷积神经网络的端到端图像压缩框架。框架包含CNN编码网络、量化、反量化、CNN解码、熵编码等几个模块，其中编解码网络可以用卷积、池化、非线性等模块进行设计和搭建。

