



# 多媒体通信 Multimedia Communications

## 第2章 多媒体数据压缩国际标准 关于图像压缩标准



2014年10月9日



下次上课时间：10月11日（星期六）上  
10月7日（星期二）的课程

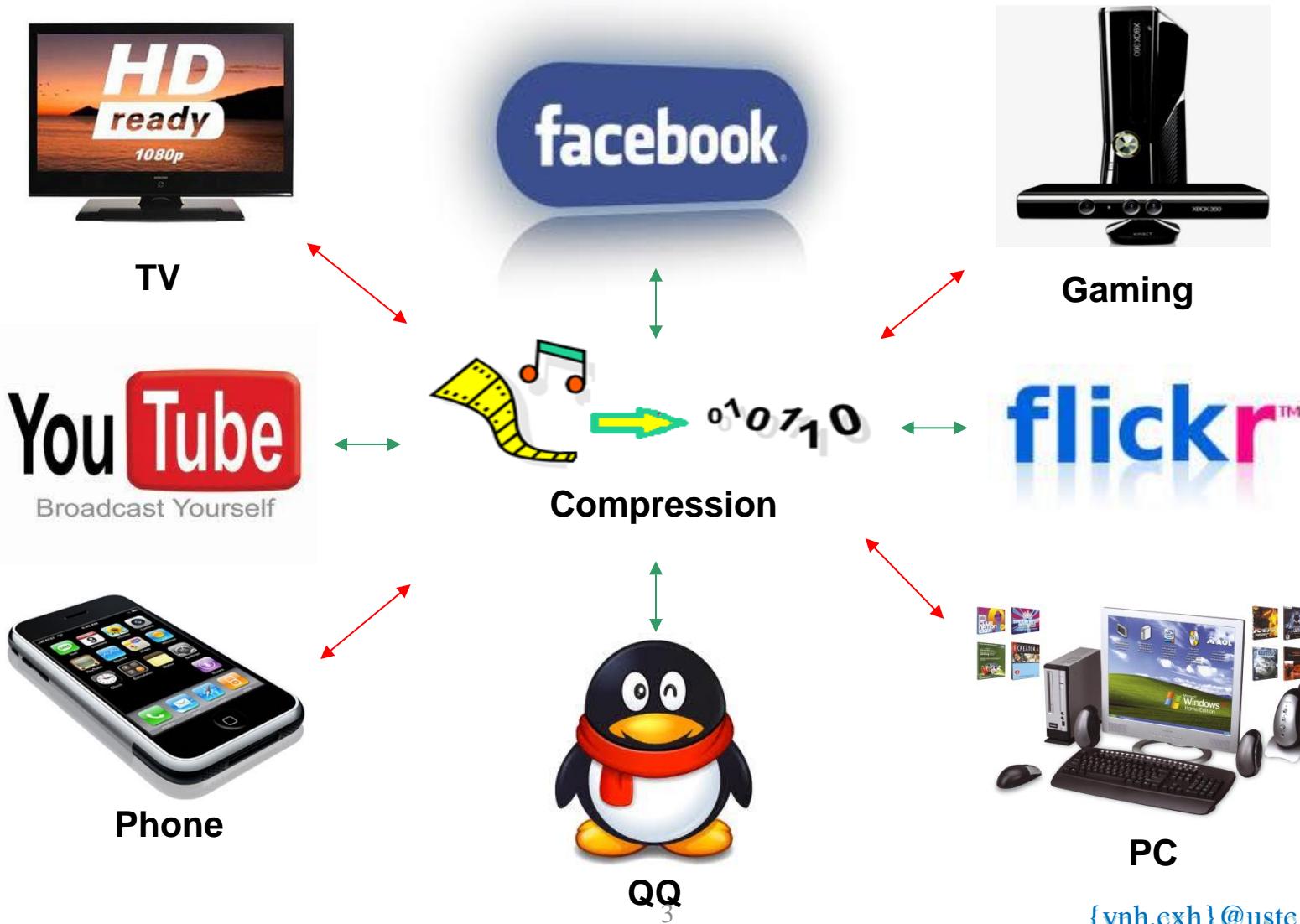


# 课程内容

- ◆ 第1章 多媒体通信概述
- ◆ 第2章 多媒体数据压缩国际标准
- ◆ 第3章 多媒体同步机制
- ◆ 第4章 多媒体网络QoS
- ◆ 第5章 多媒体传输网络
- ◆ 第6章 多媒体通信终端、系统与国际标准
- ◆ 第7章 多媒体通信技术最新进展

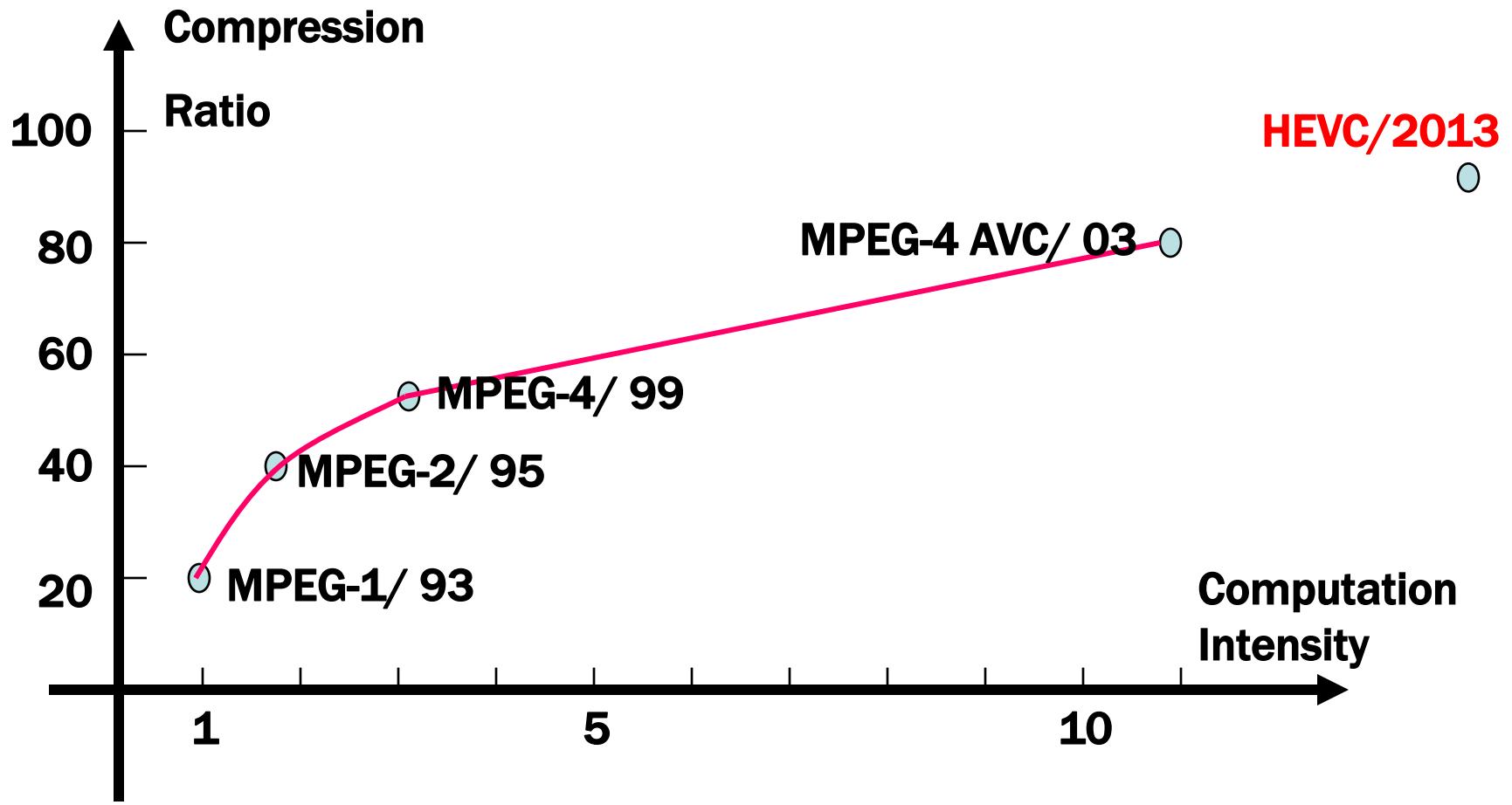


# 图像与视频的压缩





# 20年来压缩算法性能





# 20年来图像和视频压缩标准的进展

## ◆ 图像

- ◆ JPEG, 1992
- ◆ JPEG-LS(1998)
- ◆ JPEG 2000, 2000
- ◆ JPEG XR, 2009

视频压缩标准中含静止图像压缩

## ◆ 视频

- ◆ H.261, 1988
- ◆ MPEG-1, 1993
- ◆ MPEG-2/H.262, 1996
- ◆ H.263, 1996
- ◆ MPEG-4, 1999
- ◆ **MPEG-4 AVC/H.264, 2003**
- ◆ HEVC, 2013年4月13号

## ◆ China AVS, 2003





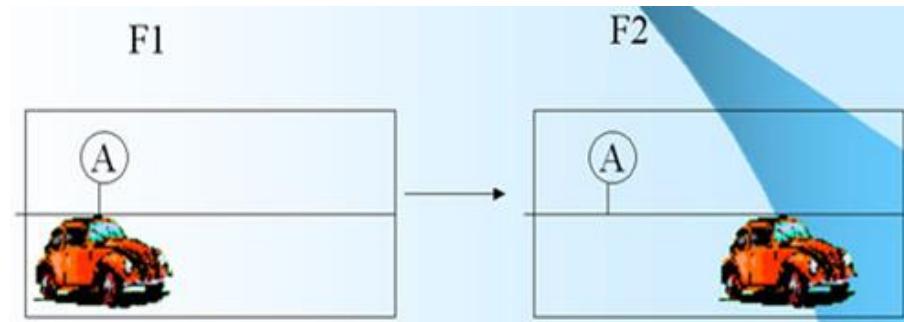
# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
  - 图像和视频的冗余
  - JPEG标准
    - JPEG标准的三个基本要素
    - JPEG中的有损和无损压缩
    - JPEG标准的工作模式
  - JPEG标准中的压缩算法
    - JPEG算法概要
    - 彩色图像 JPEG压缩
  - JPEG 2000/JPEG XR
  - H.264 Intra Frame
- ◆ 2.5 视频压缩的国际标准
- ◆ 2.6 可伸缩性编码和分布式编码



# 图像与视频的冗余：空间冗余

2幅图大块的白色背景区域构成空间上背景冗余。  
**RLE**游程编码可处理此类冗余。



图中的“**A**”是一个规则物体。光的亮度、饱和度及颜色都一样，因此，数据**A**有很大的冗余。



# 图像与视频的冗余：时间冗余

◆ 例：序列图象（动态图像）。（相邻画面的背景和移动动画面）



Frame 1

Frame 2

Frame 3

Frame 4

与PCM音频相邻样本数据间的相关性类似

{ynh,cxh}@ustc.edu.cn



# 图像与视频的冗余：熵冗余

- ◆ 源于各信息元概率的不平衡，这类冗余可以通过熵编码去除。

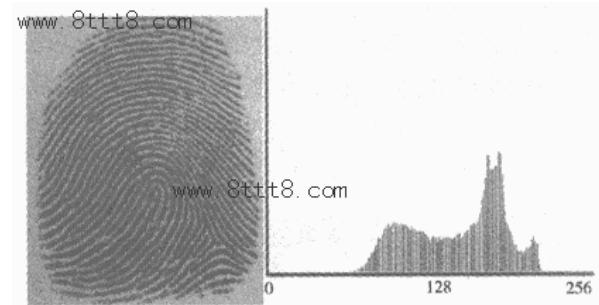
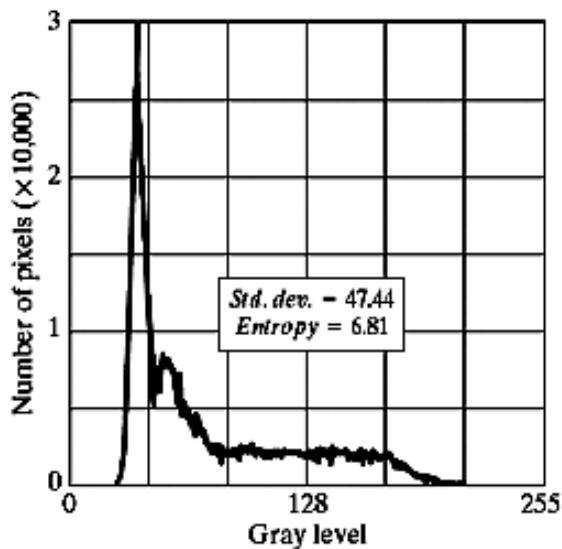


图 3 有多个波峰的指纹图像

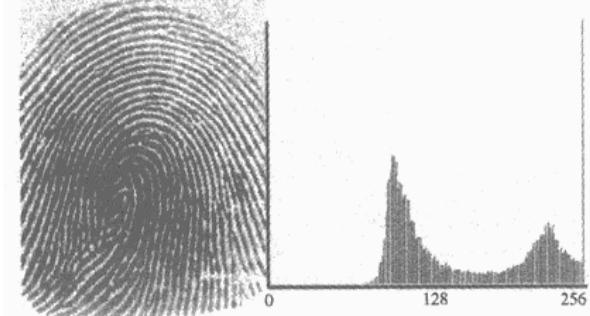


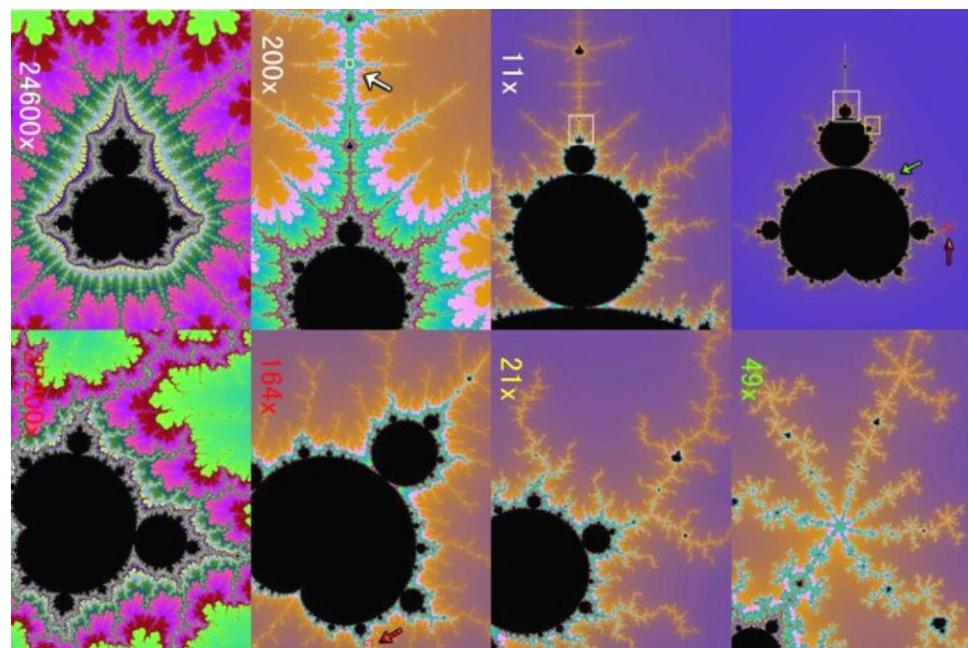
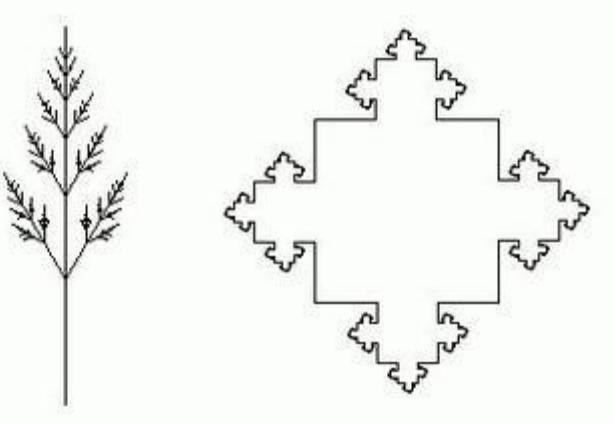
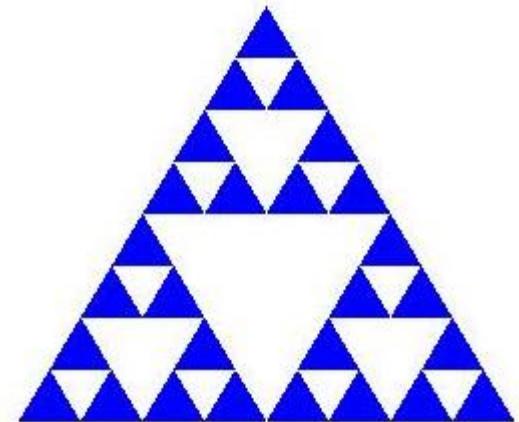
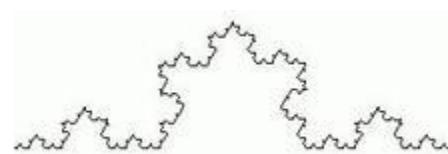
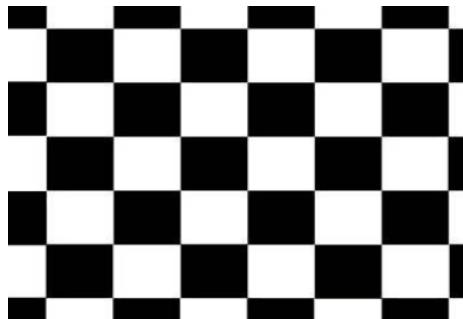
图 4 有两个波峰的指纹图像

- PCM音频样本时域数据幅度的非均匀分布
- PCM音频样本序列频率低频分量的系数大



# 图像与视频的冗余：结构冗余

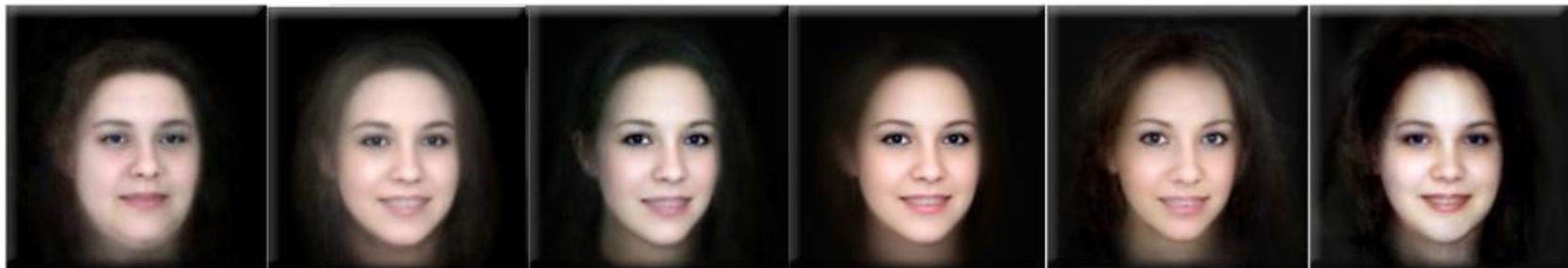
- ◆ 图象有非常强的纹理结构





# 图像与视频的冗余：知识冗余

- ◆ 有许多图像的理解与某些基础知识有相当大的相关性。例如：人脸的图像有固定的结构。嘴鼻子，眼睛等具有固定的对称和位置关系。这类规律性的结构可由先验知识相背景知识得到，我们称此类冗余为知识冗余。

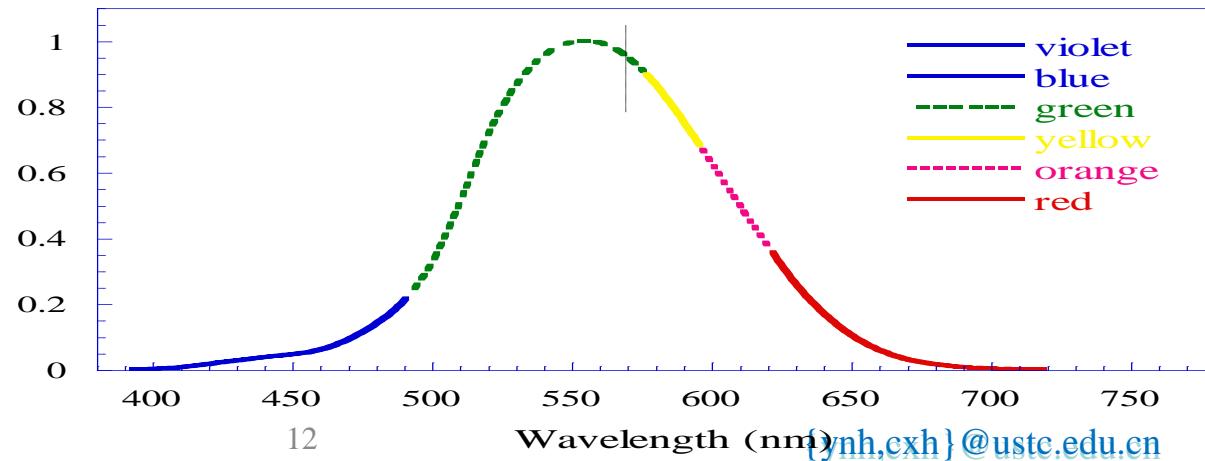
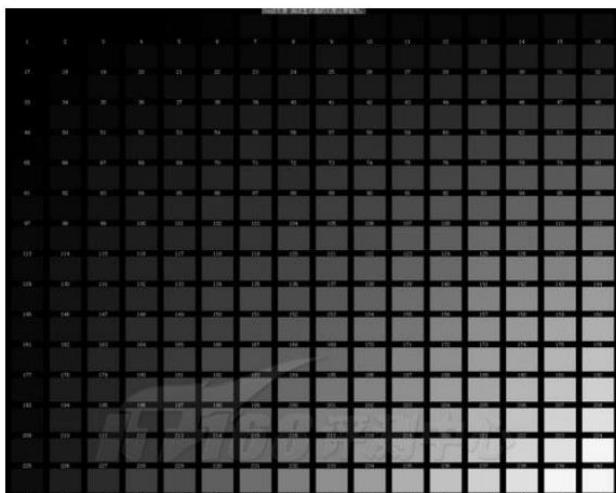




# 图像与视频的冗余：视觉冗余

◆ 人类视觉系统对于图像场某些变化是不能感知的。

- 人类视觉系统一般的分辨能力约为26灰度等级
- 对亮度变化敏感，而对色度的变化相对不敏感
- 在高亮度区，人眼对亮度变化敏感度下降
- 对物体边缘敏感，内部区域相对不敏感
- 对整体结构敏感，而对内部细节相对不敏感。





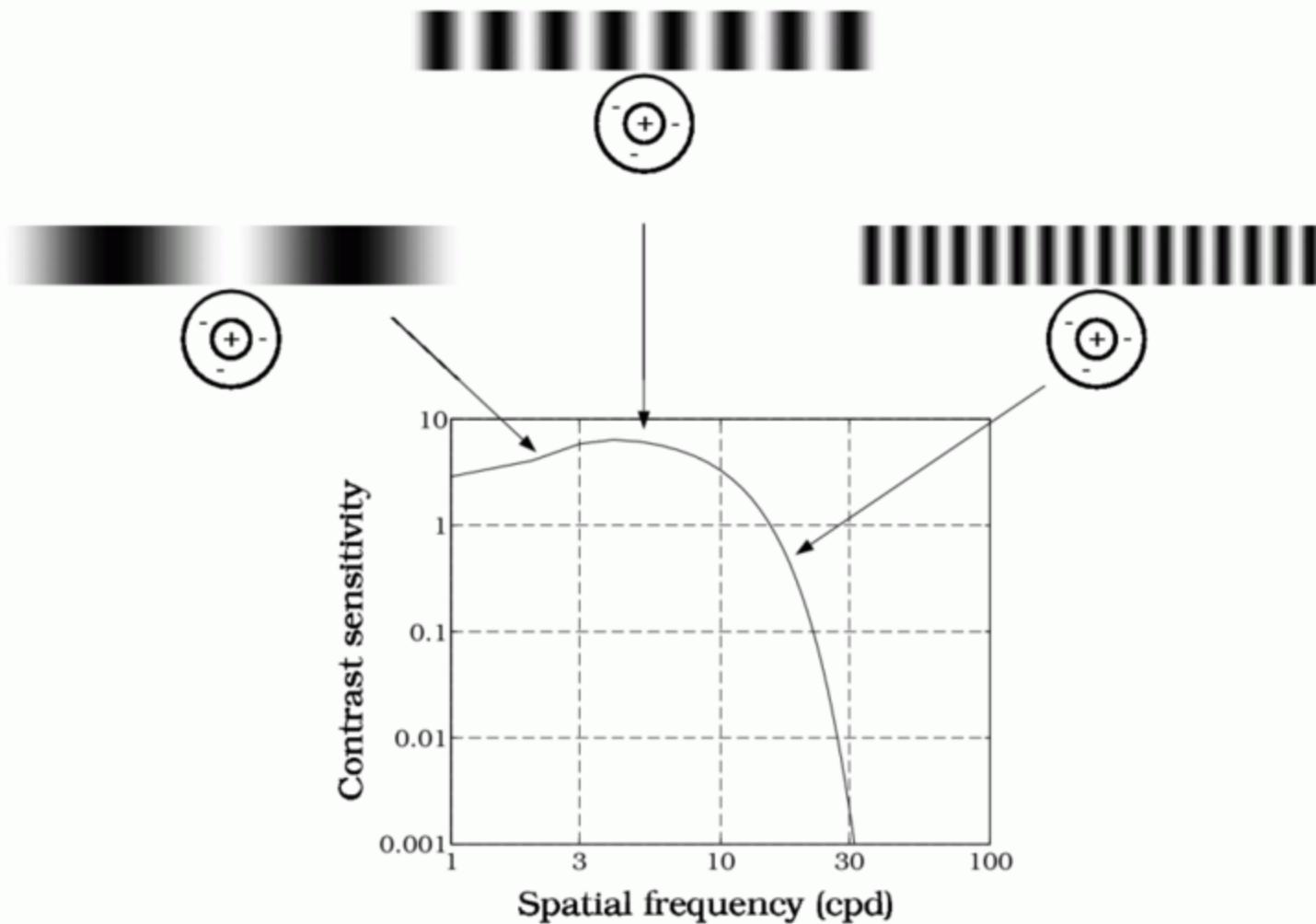
# 图像与视频的冗余：视觉冗余

## HVS：视觉敏锐度(visual acuity)

- ◆ The ability of the eye to resolve detail is known as "visual acuity." The normal human eye can distinguish patterns of alternating black and white lines with a feature size as small as one minute of an arc (**1/60 degree** or  $\pi/(60*180)$  radians).
- ◆ A few exceptional eyes may be able to distinguish features half this size. But for most of us, a pattern of higher **spatial frequency** will appear nearly pure gray.
- ◆ Low **contrast patterns** at the maximum spatial frequency will also appear gray.
- ◆ **Spatial frequency is typically measured in cycles or line pairs per millimeter (lp/mm), or cycles per degree (CPD).**



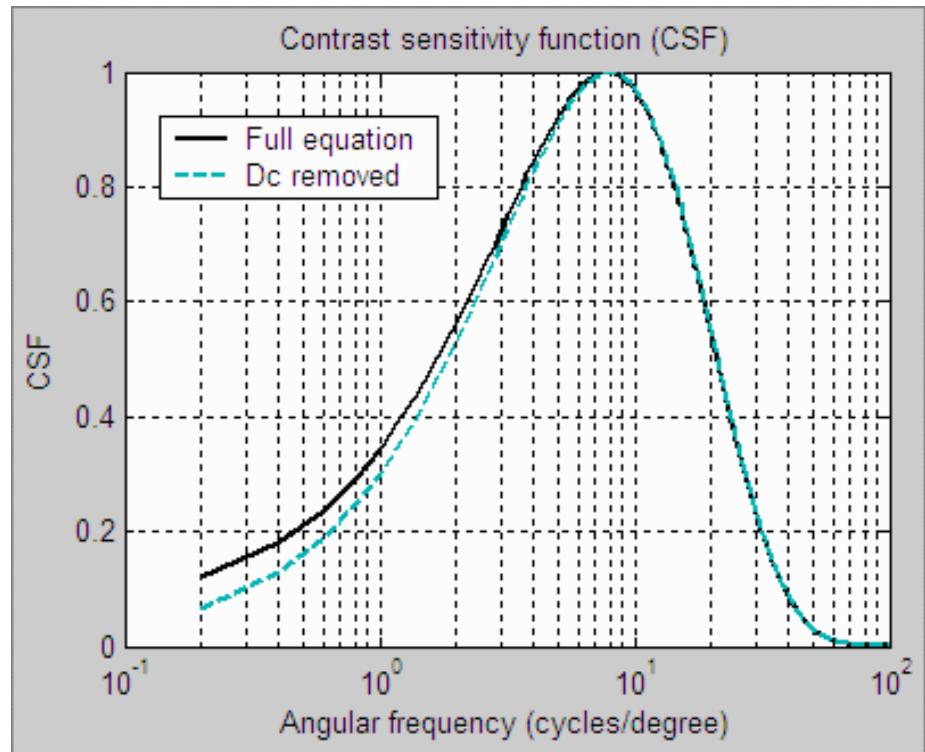
# 图像与视频的冗余：视觉冗余 contrast sensitivity vs. spatial frequency





# 对比度敏感函数 Contrast Sensitivity Function (CSF)

在空间或时间上两相邻的视觉信号，人眼刚刚能鉴别出两者差别的能力称为视觉系统的分辨率。分辨率可用**CSF**来表示。



$$\text{CSF}(f) = 2.6 (0.0192 + 0.114 f) \exp(-0.114 f)^{1.1}$$

The dc term (0.0192) can be dropped with very little effect.

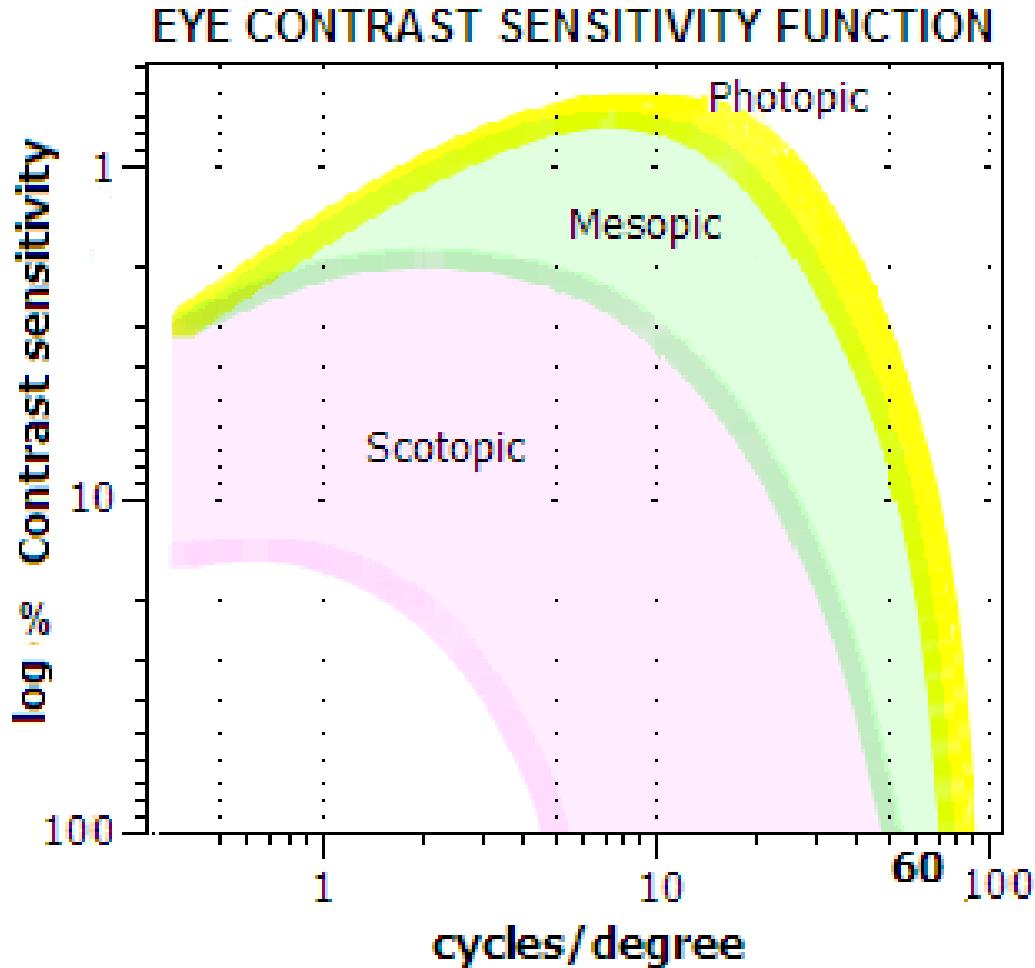
$$\text{CSF}(f) = (0.0192 + 0.114 f) \exp(-0.1254 f)$$



# 图像与视频的冗余：视觉冗余

## CSF：人眼感知区域

**Spatial frequency** is typically measured in **cycles or line pairs per millimeter** (lp/mm), which is analogous to cycles per second (Hertz) in audio systems. Lp/mm is most appropriate for film cameras, where formats are relatively fixed, but cycles/pixel (c/p) or line widths per picture height (LW/PH) may be more appropriate for digital cameras.





# 利用人眼的CSF进行压缩

- ◆ 图像系数的量化误差引起的图像变化在一定范围内是不能为人眼所觉察的。因此如果编码方案能利用HVS的一些特点，在对人眼对比度敏感的频率范围采用较小的量化步长，不够敏感的高频部分采用较大的量化步长，对亮度信号采用较小的量化步长，对色度信号采用较大的量化步长，这样是可以得到较小失真且高压缩比的图像。
- ◆ 例如，在JPEG2000中，可选择使用对不同空间频率有不同敏感度的视觉系统模型。



# 小结：图像和视频的冗余

## ◆ 从空间看冗余

- 空间冗余、熵冗余、结构冗余、知识冗余

## ◆ 从时域看冗余度

- 相邻图像之间的相关度很高

## ◆ 从频域看冗余度

- 人眼对低频分量的图像比对高频分量的图像更敏感

## ◆ 人的视觉感知机理

- 对亮度变化敏感，而对色度的变化相对不敏感
- 对物体边缘敏感，内部区域相对不敏感
- 对整体结构敏感，而对内部细节相对不敏感
- 对比度敏感函数 (CSF)



# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
  - 图像和视频的冗余
  - JPEG标准
    - JPEG标准的三个基本要素
    - JPEG中的有损和无损压缩
    - JPEG标准的工作模式
  - JPEG标准中的压缩算法
    - JPEG算法概要
    - 彩色图像 JPEG压缩
  - JPEG 2000/JPEG XR
  - H.264 Intra Frame
- ◆ 2.5 视频压缩的国际标准
- ◆ 2.6 可伸缩性编码和分布式编码



# 常见图像格式

- ◆ JPEG (Joint Photographic Experts Group)
- ◆ JPEG 2000
- ◆ Exif (Exchangeable image file format)
- ◆ TIFF (Tagged Image File Format)
- ◆ RAW refers to a family of raw image formats
- ◆ GIF (Graphics Interchange Format)
- ◆ BMP
- ◆ PNG (Portable Network Graphics)



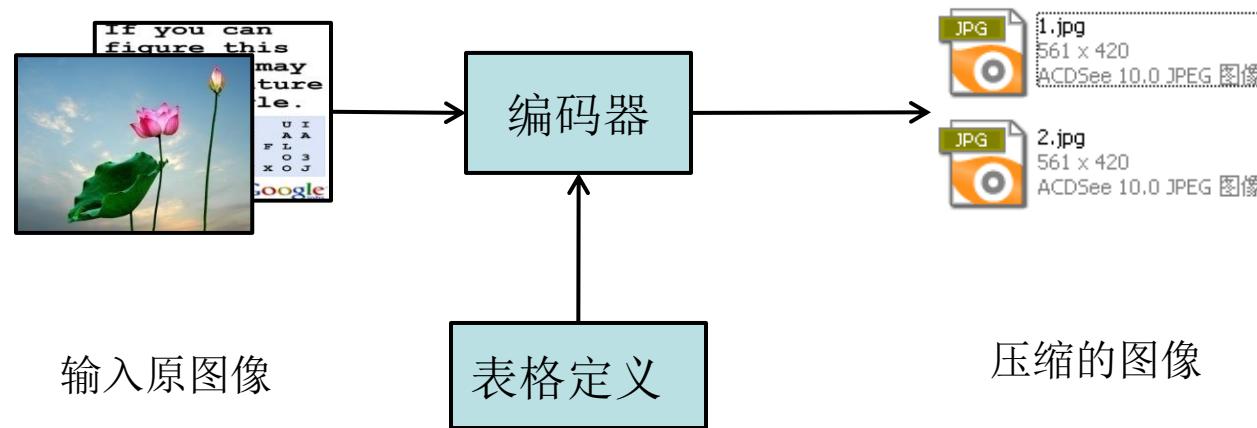
# JPEG图像压缩编码标准

- ◆ JPEG标准是“连续色调静止图象数字压缩编码”国际标准的简称。JPEG取自于联合图片专家组（Joint Photographic Expert Group）的首字母。
- ◆ JPEG工作组制订的静止图象压缩编码标准
  - JPEG(1992):规定了有损和无损(1993)两类编/解码处理
  - JPEG-LS(1998): Lossless JPEG, 无损/近无损压缩
  - JPEG2000(2000)
  - JPEG-XR(2009):JPEG extended range

# JPEG标准的三个基本要素：编码器

## ◆ 编码器

- 以合适精度将输入图像数据转换为符合交换格式的压缩图像数据
- 以合适精度将输入图像数据转换为符合简约格式的压缩图像数据

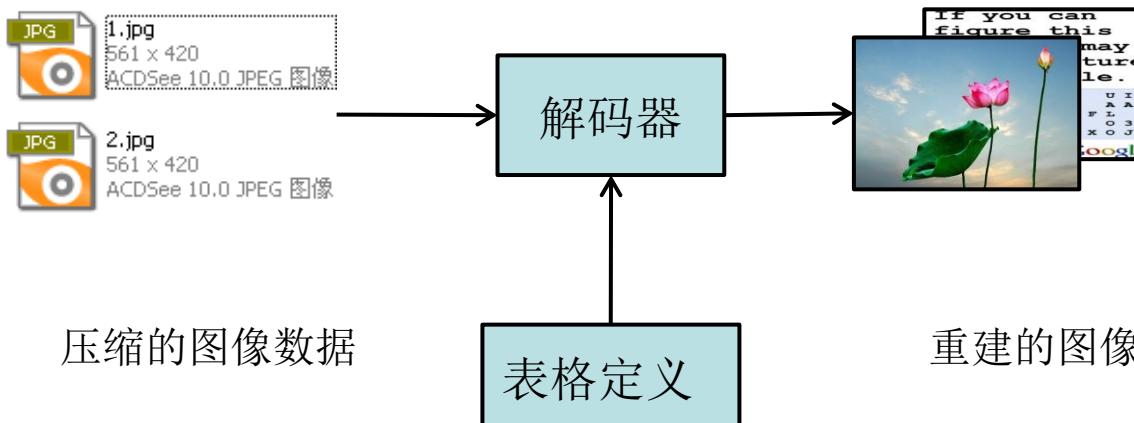




# JPEG标准的三个基本要素：解码器

## ◆ 解码器

- 接受和准确存储符合简约格式的表格数据
- 在解码器已得到解码所需的表格数据的情况下，以合适的精度由简约的压缩图像数据重建图像
- 在应用支持范围内以合适的精度将压缩的图像数据和参数转换为重建图像

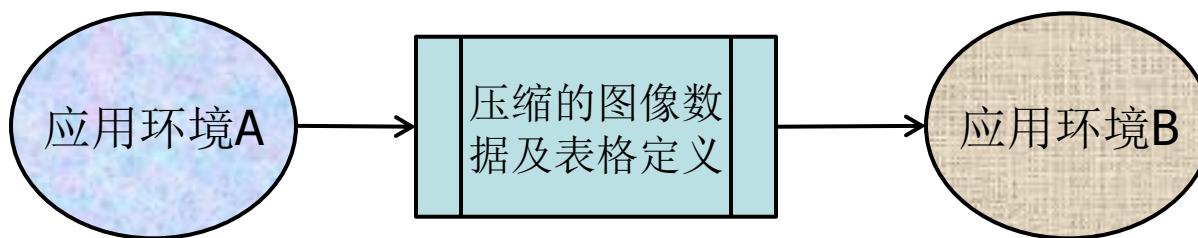




# JPEG标准的三个基本要素：交换格式 交换格式的意义

## ◆ 交换格式

- 交换格式是压缩图像数据的表示，除要求的标记字段和熵编码段外，还包括了编码中使用的所有表格。交换格式用于不同应用环境之间。





# JPEG标准的三个基本要素：交换格式 JPEG File Interchange Format, JFIF

- ◆ JPEG标准委员会没有对JPEG文件格式作出明确的定义，现在被广泛采用的是1992年9月由C-Cube Microsystems公司提出的JPEG文件交换格式（JPEG File Interchange Format, JFIF），版本号为1.02。JFIF文件格式直接使用JPEG标准为应用程序定义的许多标记，故**JFIF格式成了事实上JPEG文件交换格式标准**。
- ◆ JPEG文件大体上可以分成两个部分：**标记码**（tag）和**压缩数据**。标记码部分给出了JPEG图像的所有信息，如图像的宽、高、Huffman表、量化表等等。



# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
  - 图像和视频的冗余
  - JPEG标准
    - JPEG标准的三个基本要素
    - JPEG中的有损和无损压缩
    - JPEG标准的工作模式
  - JPEG标准中的压缩算法
    - JPEG算法概要
    - 彩色图像 JPEG压缩
  - JPEG 2000/JPEG XR
  - H.264 Intra Frame
- ◆ 2.5 视频压缩的国际标准
- ◆ 2.6 可伸缩性编码和分布式编码



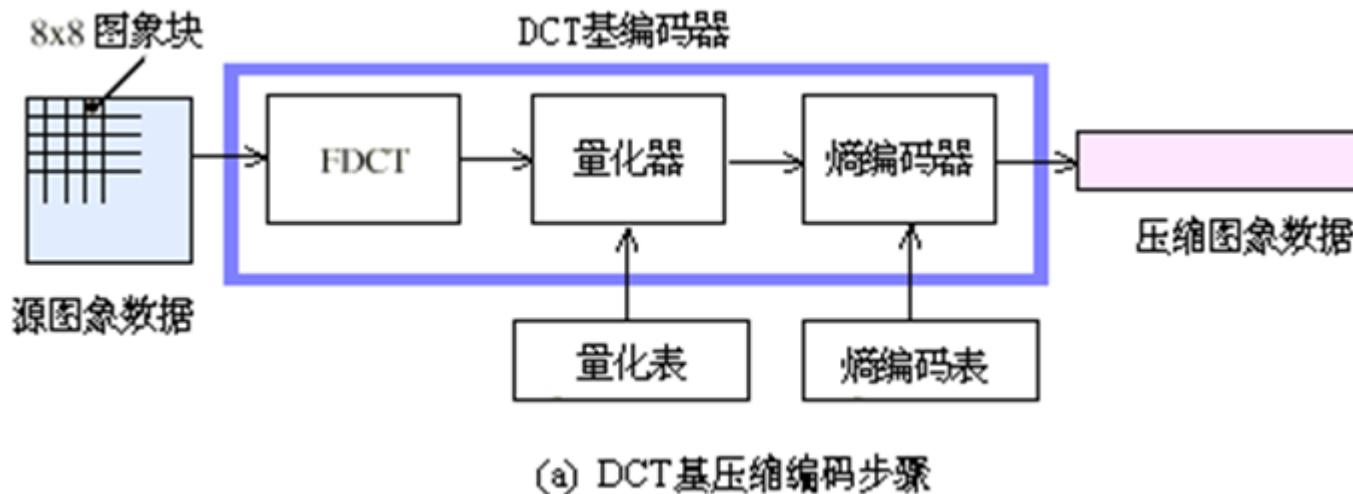
# JPEG中的有损和无损压缩

- ◆ JPEG规定了有损和无损两类编码和解码处理
  - 基于离散余弦变换**DCT**的编码，它提供有损压缩，能在保证解码重建图像的视觉逼真度的条件下，达到对图像数据的显著压缩；
  - 基于**二维预测**的编码，它提供无损压缩。能在保持图像质量的情况下尽可能的压缩图像数据；



# JPEG中的有损压缩： 基于DCT的编码

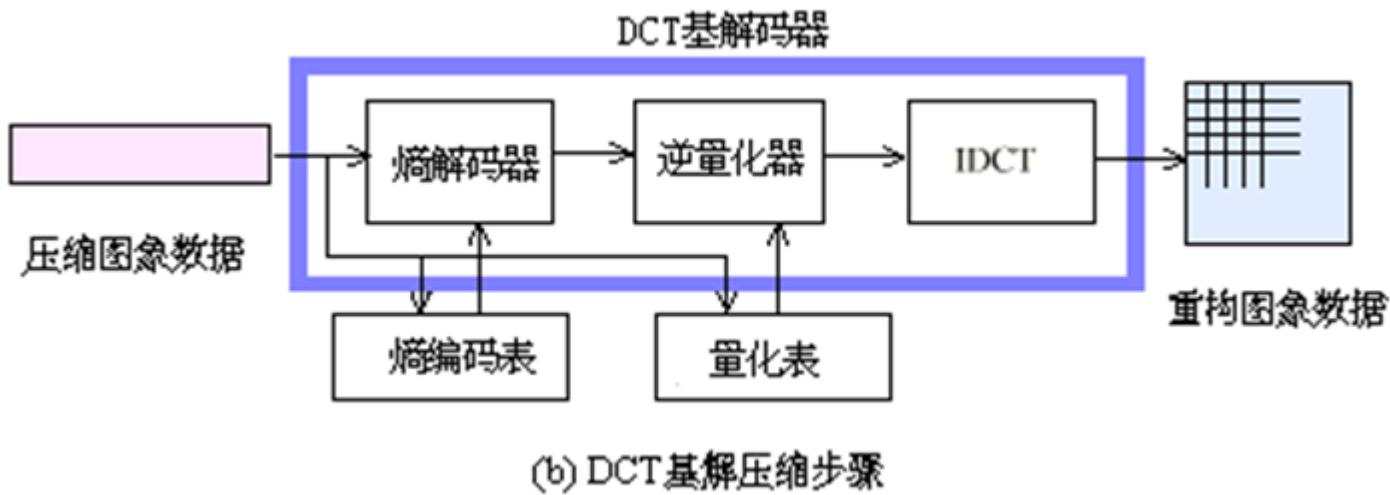
- ◆ 图像分量被分割成不重叠的8\*8小块，每个子块通过DCT（FDCT）变换成64个DCT系数（1个直流DC系数和63个交流AC系数），然后利用量化表对系数进行均匀量化，对量化系数进行处理（DC预测和Zig-zag扫描）后，进行熵编码（霍夫曼编码或者算术编码）。





# JPEG中的有损压缩： 基于DCT的解码

- ◆ 解码是编码的逆过程，每一个编码单元在解码端都有一个相应的逆处理单元与之对应，各单元的排列顺序也正好相反

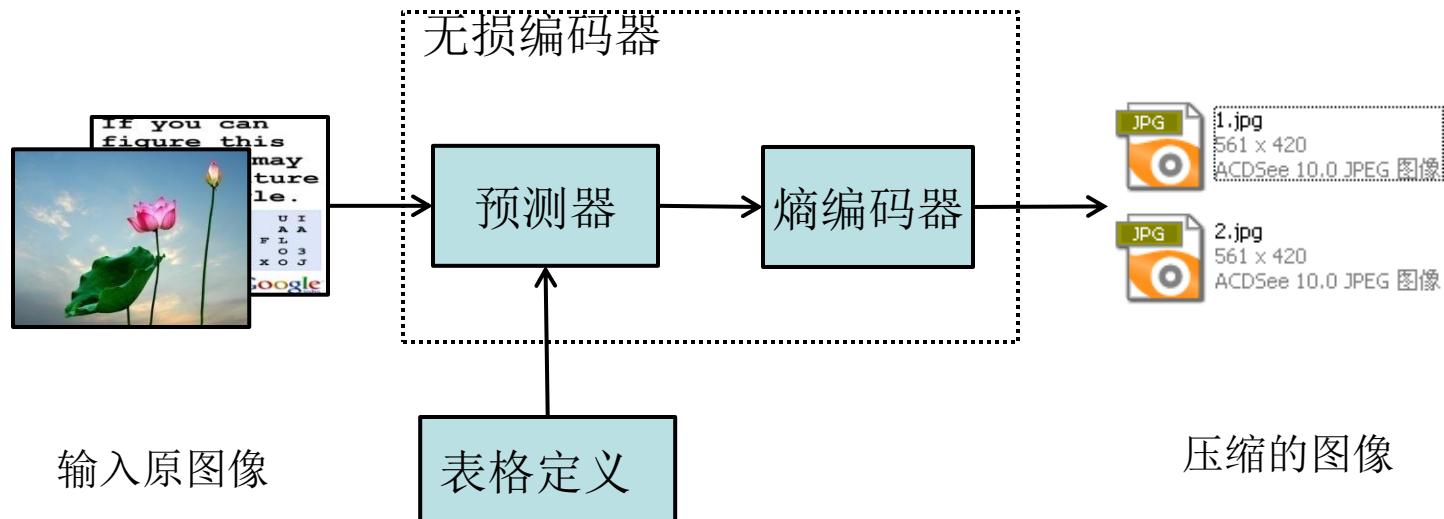




# JPEG中的无损压缩： 基于二维预测的无损编码

## ◆ 无损编码处理的主要步骤

- 无损编码先对各像素进行**3邻域预测编码**，然后对预测误差进行霍夫曼编码或者算术编码





# JPEG中的无损压缩： 基于二维预测的无损编码

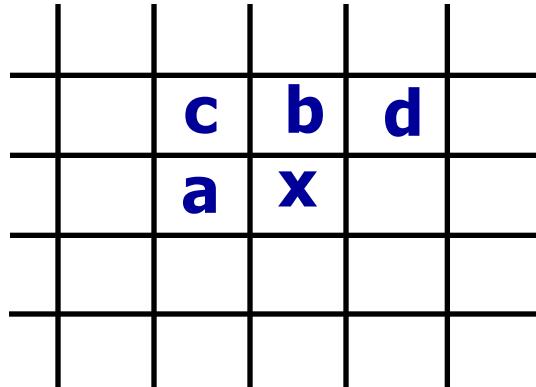
- ◆ JPEG提供了两种可供选择的熵编码：霍夫曼编码和算术编码。

| 编码方法  | 码表                | 应用              |
|-------|-------------------|-----------------|
| 霍夫曼编码 | 霍夫曼码表<br>无缺省，必须指定 | 所有四种模式          |
| 算术编码  | 状态表<br>有缺省选择      | 除基本系统外的<br>三种模式 |



# JPEG中的无损压缩： 静止图像的二维预测编码原理

- ◆这种压缩算法被应用到JPEG标准的无损压缩模式之中，中等复杂程度的图像压缩比可达到**2:1**。



三邻域预测法

| 选择值 | 预测值         |
|-----|-------------|
| 0   | 非预测         |
| 1   | a           |
| 2   | b           |
| 3   | c           |
| 4   | $a+b-c$     |
| 5   | $a+(b-c)/2$ |
| 6   | $b+(a-c)/2$ |
| 7   | $(a+b)/2$   |

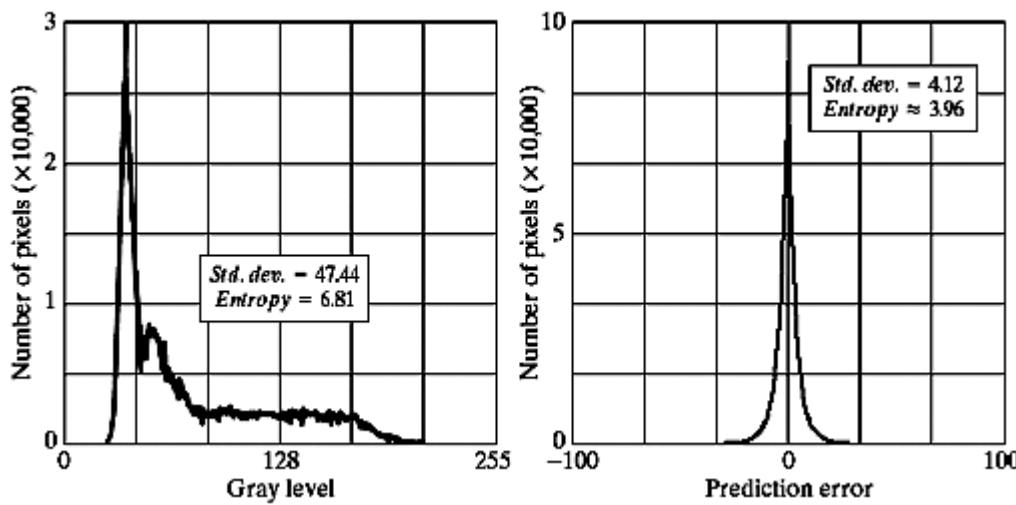


# JPEG中的无损压缩： 静止图像的二维预测编码效果示例

原图



预测残差  
(用128表示0)



亮度直方图



# JPEG中的无损压缩： 静止图像无损编码的其他方式

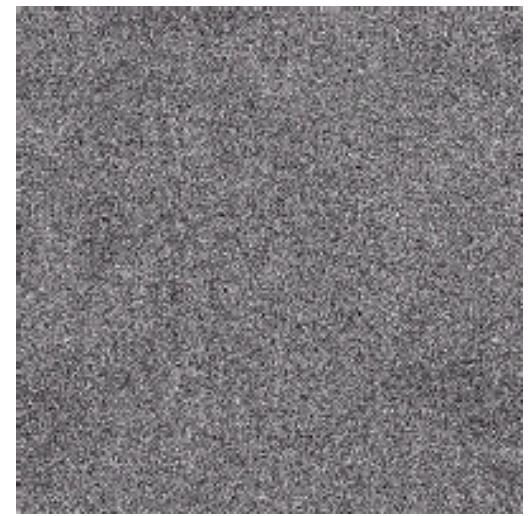
- ◆ 无损编码也可以稍加改变的使用，方法是将输入图像采样值先**减少一个或者多个比特**，然后再进行无损压缩编码，这比正常无损编码有更高的压缩率。在同样的视觉逼真度下，压缩率低于基于DCT的处理，但是却能限制重建图像的最大误差



Original 8bits/pixel



Bitplane 7



Bitplane 0

{ynh,cxh}@ustc.edu.cn

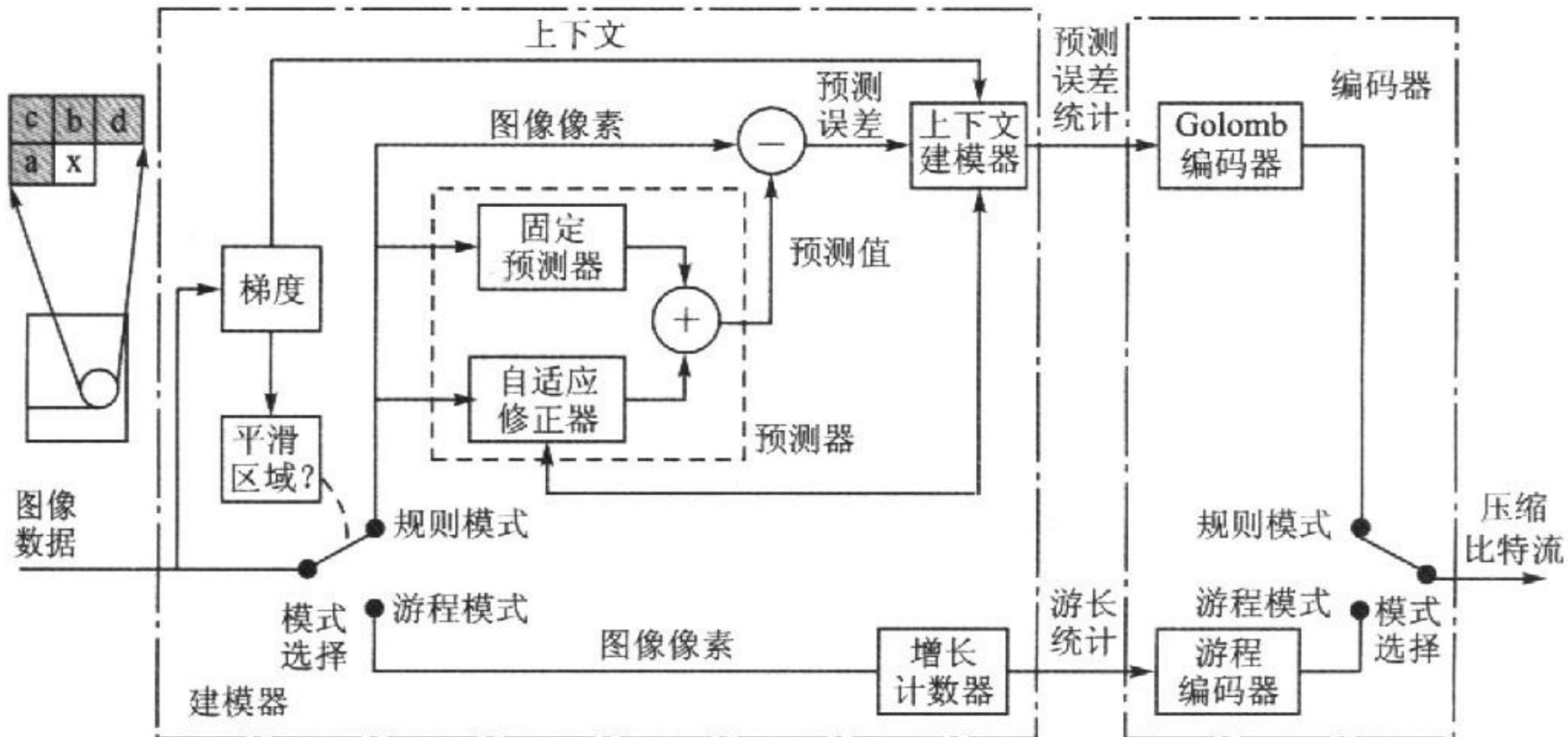


# JPEG-LS(1998)

- ◆ JPEG-LS is the new lossless/**near-lossless** compression standard for continuous-tone images, ISO-14495-1/ITU-T.87. The standard is based on the LOCO-I algorithm (LOw COMplexity LOssless COmpression for Images) developed at Hewlett-Packard Laboratories.
- ◆ Compression for JPEG-LS is generally much **faster than JPEG 2000** and much better than the original lossless **JPEG** standard.



# JPEG-LS编码流程



- 三邻域预测结果与边缘检测结果相关
- 根据梯度计算是否为平滑区域（平滑区域熵编码效果不好）



# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
  - 图像和视频的冗余
  - JPEG标准
    - JPEG标准的三个基本要素
    - JPEG中的有损和无损压缩
    - **JPEG标准的工作模式**
  - JPEG标准中的压缩算法
    - JPEG算法概要
    - 彩色图像 JPEG压缩
  - JPEG 2000/JPEG XR
  - H.264 Intra Frame
- ◆ 2.5 视频压缩的国际标准
- ◆ 2.6 可伸缩性编码和分布式编码



# JPEG标准的工作模式

- ◆ JPEG标准定义了四种不同的工作模式，分别为
  - 无损工作模式：如前所述
  - 基于DCT的顺序处理模式
  - 基于DCT的渐进处理模式
  - 分等级处理模式



## JPEG标准的工作模式 基于DCT的顺序处理模式（JPEG基本系统）

- ◆ 输入原图像按从上至下，从左至右的次序被分成 $8*8$ 的不重叠子块，当一个子块经过前向DCT变换、量化和针对熵编码的转换之后（DC系数预测和Zig-zag扫描），立即进行熵编码，并称为输出图像的一部分。这种处理过程仅需一个子块的数据存储空间。



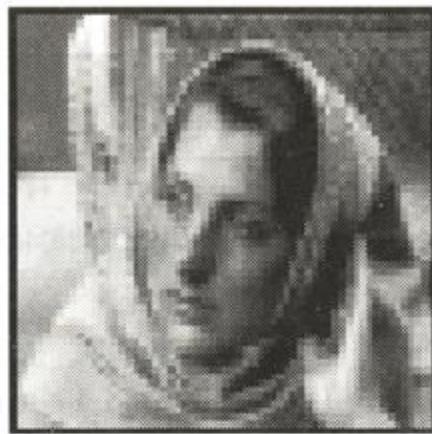
# JPEG标准的工作模式

## 基于DCT的渐进处理模式

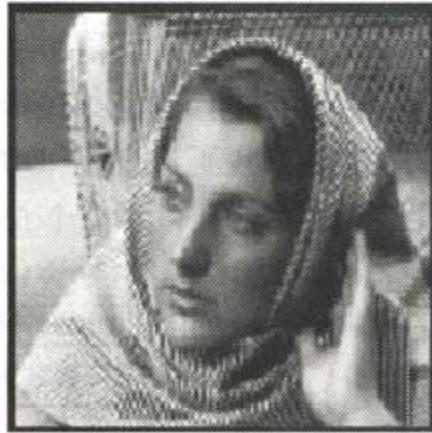
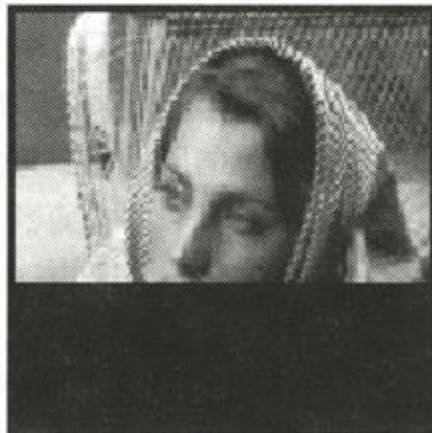
- ◆ 子块按同样的次序进行编码，但整个图像要进行若干次扫描完成编码。在量化器和熵编码器之间有一个与图像大小相同的系数贮存器。当每一个字块进行DCT变换和量化之后，量化系数存入该贮存器内。然后在每一次扫描时将这些系数的一部分进行编码。
- ◆ 对贮存的DCT系数，有两种不同的扫描方式：
  - 频谱选择法：在之字形扫描系数时，只选择部分指定的频率域变换系数进行编码。通常是在前面的扫描先对低频分量进行编码，然后在之后的扫描对更高的频率分量进行编码，直到所有的系数编码完毕
  - 连续近似法：对DCT系数按位平面从高到低进行扫描。在前面的扫描先将最高几个有效比特的值进行编码，后续扫描对剩下的比特位进行编码
- ◆ 渐进模式中，两种方法可分别使用，也可灵活的组合使用



# JPEG标准的工作模式



(a) 渐进显示



(b) 顺序显示

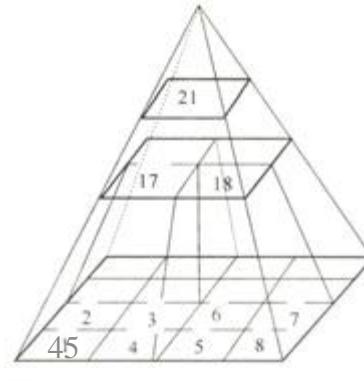


# JPEG标准的工作模式

## 分等级处理模式

◆ 图像被当作一个帧系列进行编码。这些帧提供了重建分量的参考帧并用于后续帧的预测。除第一帧外使用帧间差进行编码(基于DCT的或者无损编码抑或组合使用)。编码中可以使用下抽样和上抽样滤波，已得到金字塔状的不同空间分辨率。此外在给定分辨率的条件下，分等级编码模式也可用于改进重建图像分量的质量。

◆ 分等级编码模式提供了与渐进DCT编码模式类似的逐渐显示功能，但在要求多分辨率的应用场合更为有用（不同解码和显示能力的终端可以通用）。分等级模式也提供了渐进的最终无损编码的能力。



(a) 一幅图像的多分辨率金字塔表示

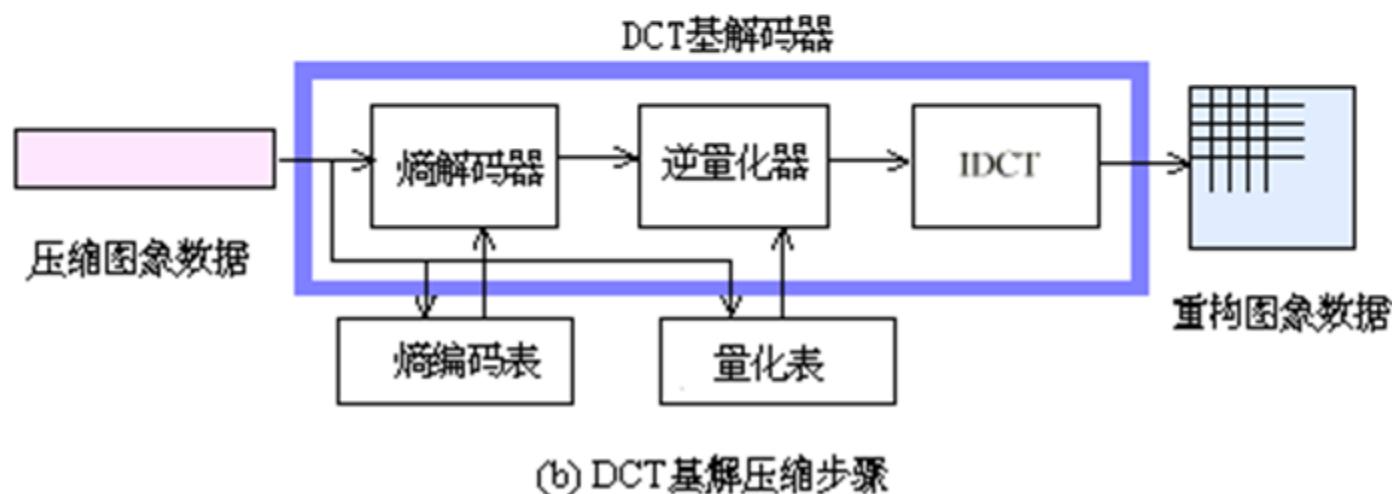
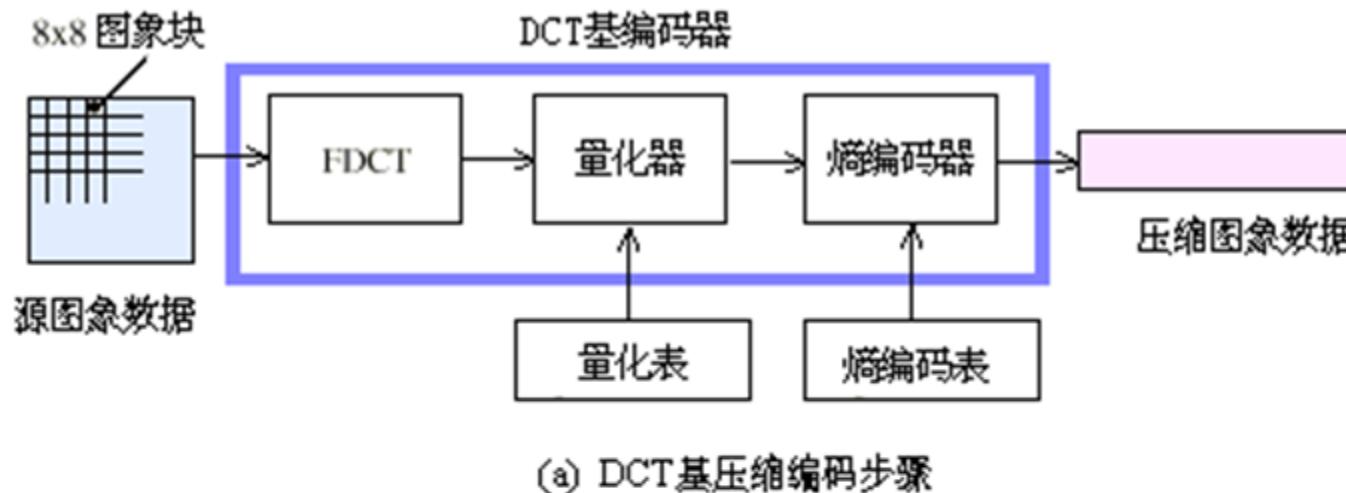


# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
  - 图像和视频的冗余
  - JPEG标准
    - JPEG标准的三个基本要素
    - JPEG中的有损和无损压缩
    - JPEG标准的工作模式
  - JPEG标准中的压缩算法
    - **JPEG算法概要**
    - 彩色图像 JPEG压缩
  - JPEG 2000/JPEG XR
  - H.264 Intra Frame
- ◆ 2.5 视频压缩的国际标准
- ◆ 2.6 可伸缩性编码和分布式编码



# JPEG算法基本步骤





# JPEG算法概要

- ◆ 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. 使用RLE对交流系数(AC)进行编码
- ◆ 6. 熵编码(entropy coding)

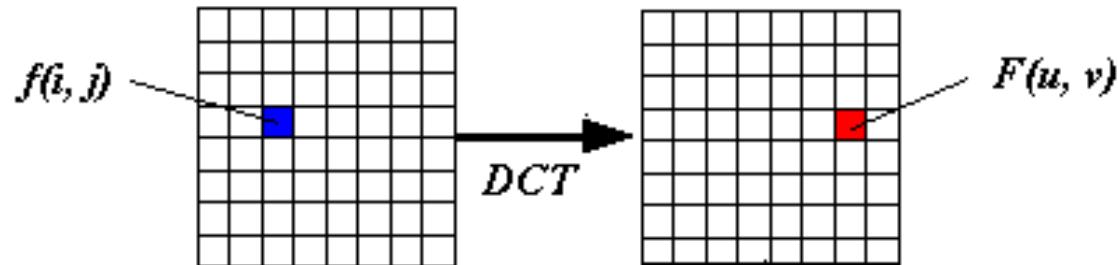


# 正向离散余弦变换(FDCT)

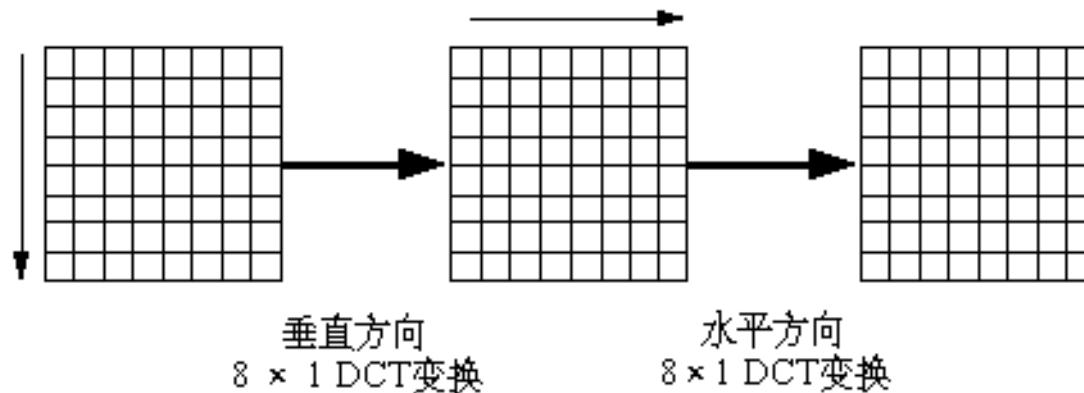
- ◆ 对每个单独的彩色图像分量，把整个分量图像分成 $8 \times 8$ 的图像块，作为两维离散余弦变换DCT的输入。通过DCT变换，把能量集中在少数几个系数上。
- ◆ 通常在进行DCT变换前，先将像素值进行幅度平移，使其成为一个有符号数。方法是减去 $2^{p-1}$ ， $p$ 为像素位深度。对于普通256级灰度图像，即减去128。这样能减少图像能量，变换后系数更小。



# DCT计算



- ◆ 在计算二维的DCT变换时，可把二维的DCT变换变成一维的DCT变换

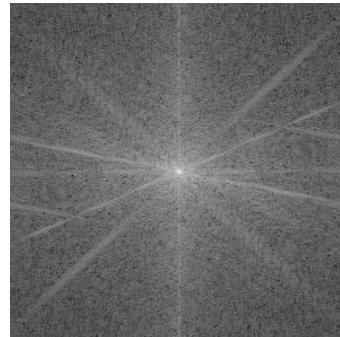




# DCT变换



original



DFT



DCT

## Transform selection

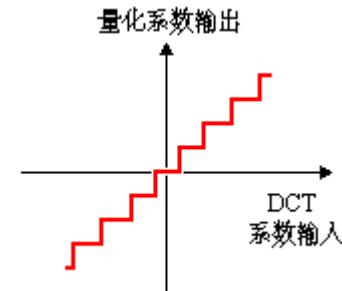
Information packing ability:  $KLT > DCT > DFT$

Computational complexity:  $DCT < DFT < KLT$



# 量化

- ◆ 量化是对经过FDCT变换后的频率系数进行量化。量化的目的是减小非“0”系数的幅度以及增加“0”值系数的数目。**量化是图像质量下降的最主要原因。**
- ◆ 对于有损压缩算法，JPEG算法使用均匀量化器进行量化，量化步距是按照系数所在的位置和每种颜色分量的色调值来确定。因为**人眼对亮度信号比对色差信号更敏感**，因此使用了两种量化表。





# 量化表

- ◆ 由于人眼对低频分量的图像比对高频分量的图像更敏感，因此图中的左上角的量化步距要比右下角的量化步距小。
- ◆ 表中的数值对CCIR 601标准电视图像已经是最佳的。还可以使用自己的量化表。

|    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|
| 17 | 18 | 24 | 47 | 99 | 99 | 99 | 99 | 99 |
| 18 | 21 | 26 | 66 | 99 | 99 | 99 | 99 | 99 |
| 24 | 26 | 56 | 99 | 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 | 99 | 99 | 99 | 99 | 99 |

|    |    |    |    |     |     |     |     |
|----|----|----|----|-----|-----|-----|-----|
| 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  |

哪一个是亮度量化表？



## Quantization Table Used

|    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|
| 8  | 5  | 5  | 8  | 12 | 20 | 25 | 30 |
| 6  | 6  | 7  | 9  | 13 | 29 | 30 | 27 |
| 7  | 6  | 8  | 12 | 20 | 28 | 34 | 28 |
| 7  | 8  | 11 | 14 | 25 | 43 | 40 | 31 |
| 9  | 11 | 18 | 28 | 34 | 54 | 51 | 38 |
| 12 | 17 | 27 | 32 | 40 | 52 | 56 | 46 |
| 24 | 32 | 39 | 43 | 51 | 60 | 60 | 50 |
| 36 | 46 | 47 | 49 | 56 | 50 | 51 | 49 |

## Compressed Image



Compression Ratio: 7.7

|    |    |    |    |     |     |     |     |
|----|----|----|----|-----|-----|-----|-----|
| 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  |



Compression Ratio: 12.3

|     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|
| 64  | 44  | 40  | 64  | 96  | 160 | 204 | 244 |
| 48  | 48  | 56  | 76  | 104 | 232 | 240 | 220 |
| 56  | 52  | 64  | 96  | 160 | 228 | 276 | 224 |
| 56  | 68  | 88  | 116 | 204 | 300 | 300 | 248 |
| 72  | 88  | 148 | 224 | 272 | 300 | 300 | 300 |
| 96  | 140 | 220 | 256 | 300 | 300 | 300 | 300 |
| 196 | 256 | 300 | 300 | 300 | 300 | 300 | 300 |
| 288 | 300 | 300 | 300 | 300 | 300 | 300 | 300 |



Compression Ratio: 33.9

|     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|
| 128 | 128 | 128 | 128 | 128 | 128 | 128 | 128 |
| 128 | 128 | 128 | 128 | 128 | 128 | 128 | 128 |
| 128 | 128 | 128 | 128 | 128 | 128 | 128 | 128 |
| 128 | 128 | 128 | 128 | 128 | 128 | 128 | 128 |
| 128 | 128 | 128 | 128 | 128 | 128 | 128 | 128 |
| 128 | 128 | 128 | 128 | 128 | 128 | 128 | 128 |
| 128 | 128 | 128 | 128 | 128 | 128 | 128 | 128 |
| 128 | 128 | 128 | 128 | 128 | 128 | 128 | 128 |



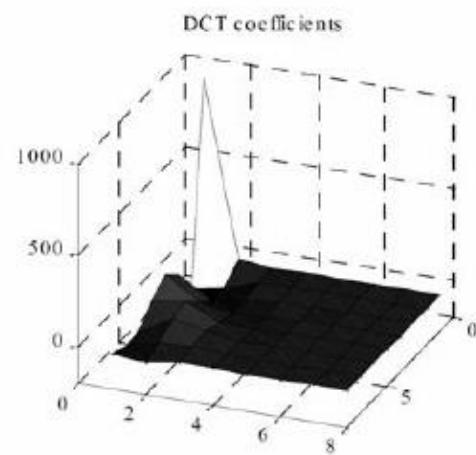
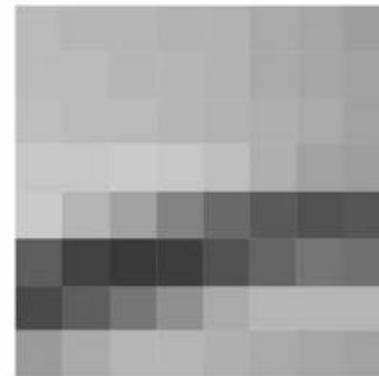
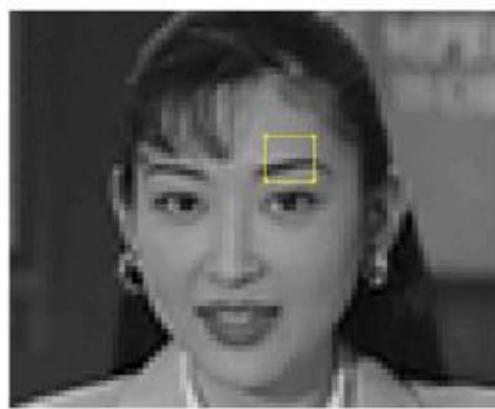
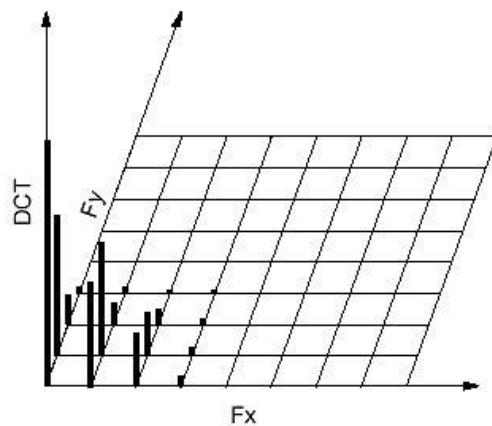
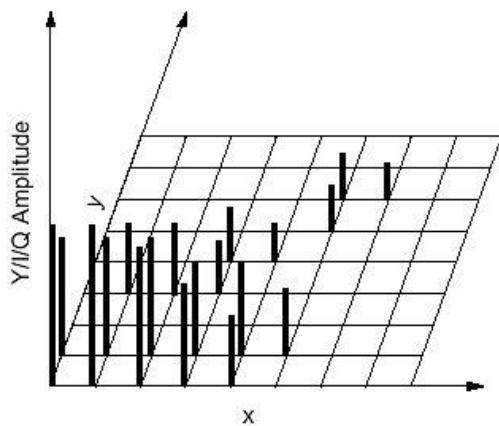
Compression Ratio: 60.1

# JPEG Example



Original  
Image

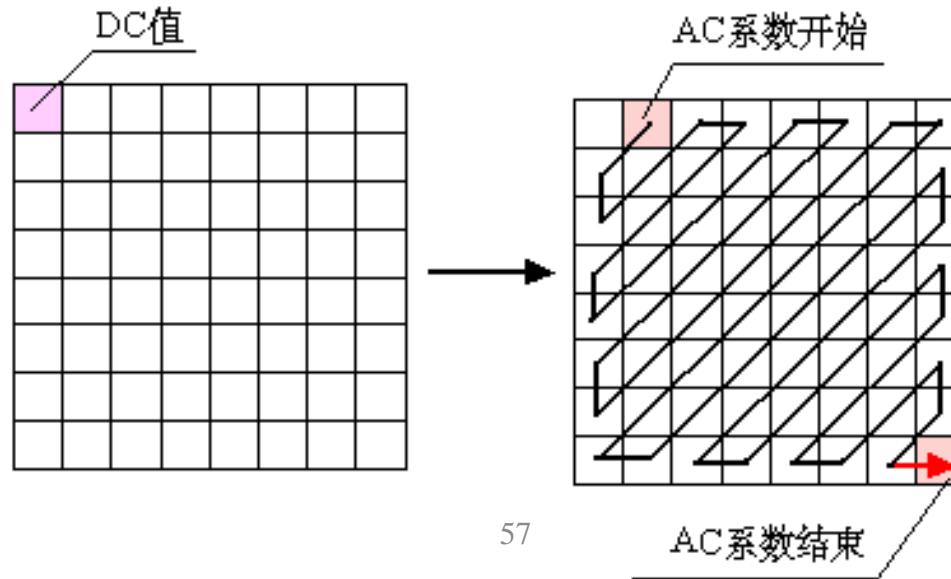
# DCT结果示例





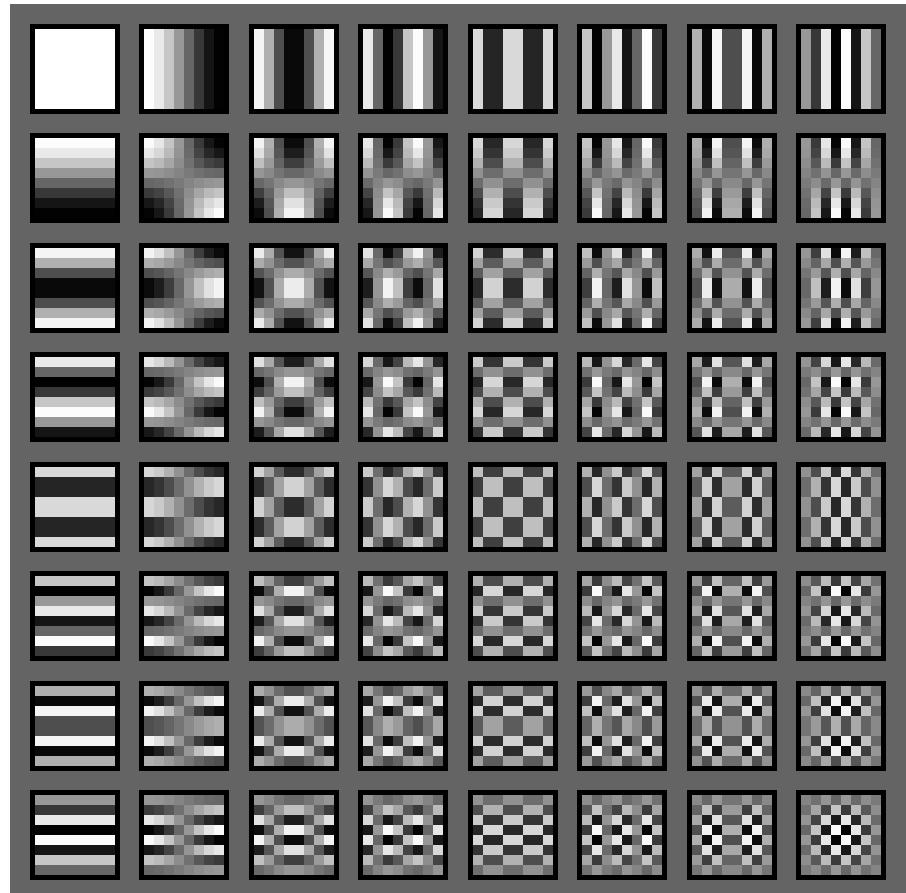
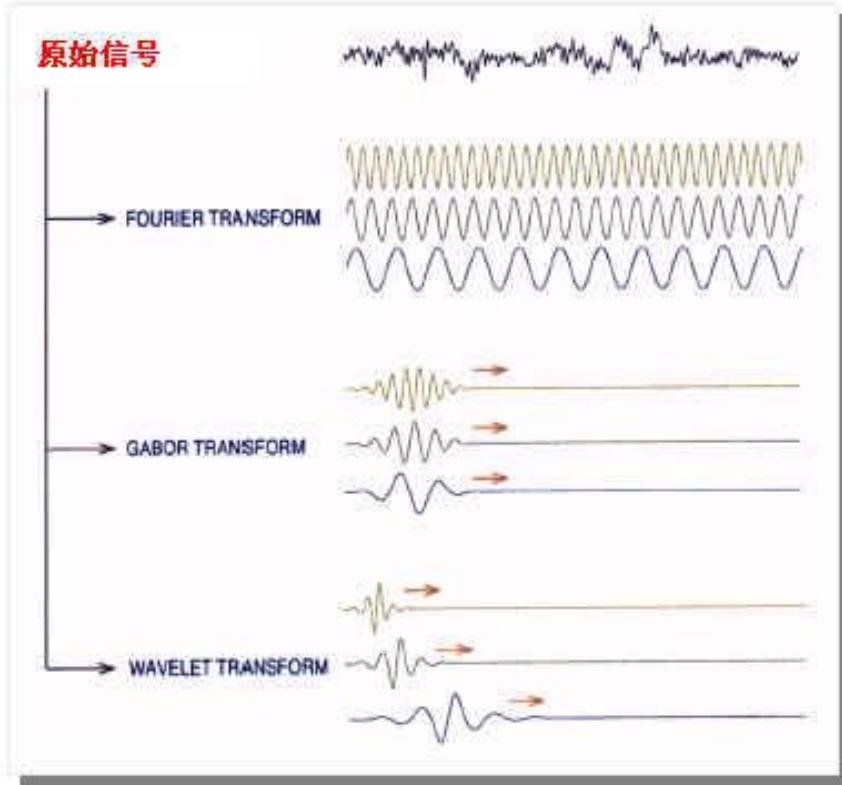
# Zig-zag

◆量化后的系数要重新编排，目的是为了增加连续的“0”系数的个数，就是“0”的游程长度，方法是按照Z字形的式样编排。这样就把一个 $8 \times 8$ 的矩阵变成一个 $1 \times 64$ 的矢量，频率较低的系数放在矢量的顶部。





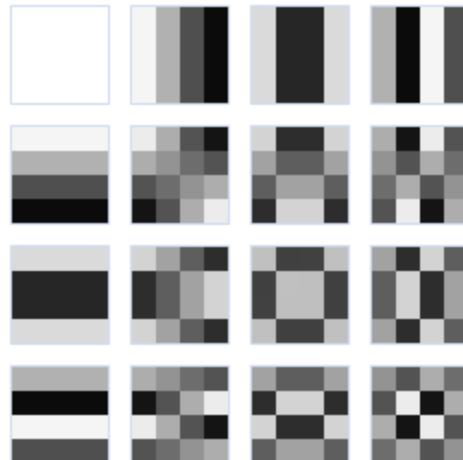
# 64 (8 x 8) DCT basis functions



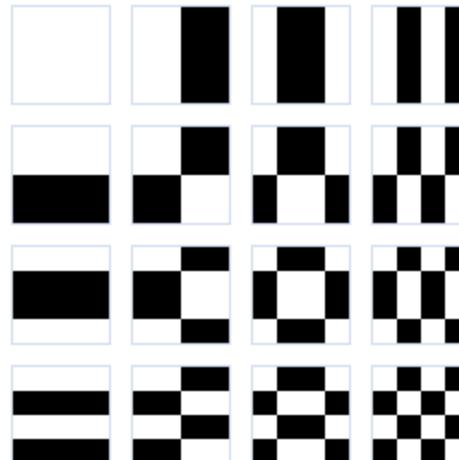


# 4x4 blocks 不同变换的基函数

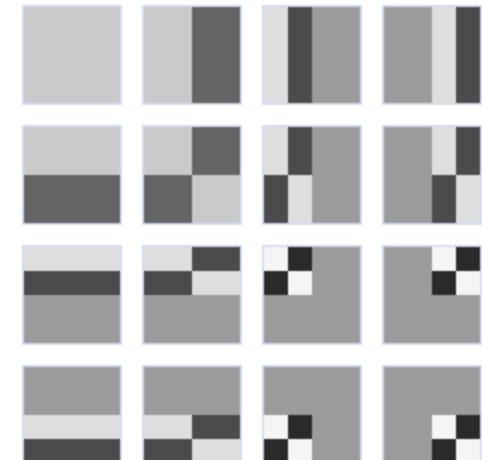
- ◆ 4x4 Discrete Cosine Transform
- ◆ 4x4 Walsh-Hadamard Transform (WHT)
- ◆ 4x4 Haar Wavelet Transform



DCT



Walsh-Hadamard



Haar Wavelet



# Gabor变换的基函数

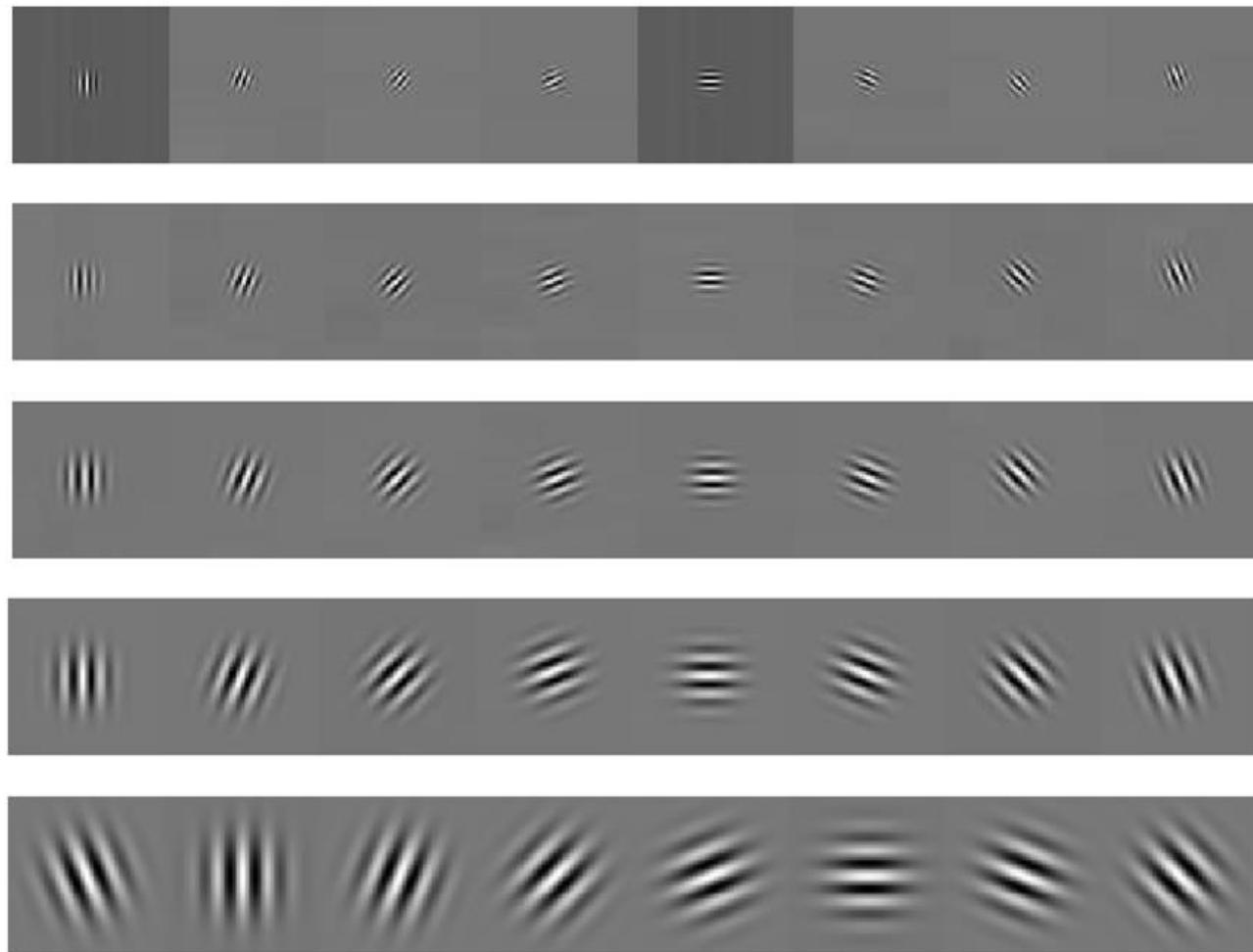


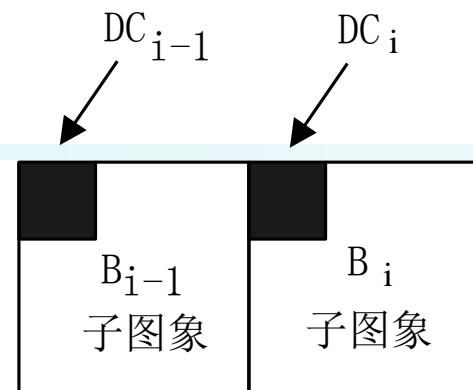
图 4.1 5 个尺度 8 个方向的 Gabor 滤波器核函数示意图



# 量化后的编码

- ◆ 直流系数的编码：DPCM
- ◆ 交流系数的编码：RLE
- ◆ 熵编码

□ 使用熵编码还可以对DPCM编码后的直流DC系数和RLE编码后的交流AC系数作进一步的压缩。在JPEG有损压缩算法中，使用霍夫曼编码器来减少熵。使用霍夫曼编码器的理由是可以使用很简单的查表(lookup table)方法进行编码。



$$DC\text{差分值} = DC_i - DC_{i-1}$$



# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
  - 图像和视频的冗余
  - JPEG标准
    - JPEG标准的三个基本要素
    - JPEG中的有损和无损压缩
    - JPEG标准的工作模式
  - JPEG标准中的压缩算法
    - JPEG算法概要
    - 彩色图像 JPEG压缩
  - JPEG 2000/JPEG XR
  - H.264 Intra Frame
- ◆ 2.5 视频压缩的国际标准
- ◆ 2.6 可伸缩性编码和分布式编码



# 彩色图像 JPEG压缩流程

◆JPEG文件使用的颜色空间是电视图像信号数字化标准ITU-RBT 601推荐标准规定的YC<sub>b</sub>C<sub>r</sub>彩色空间。从**RGB**转换成**YCbCr**的计算公式如下：

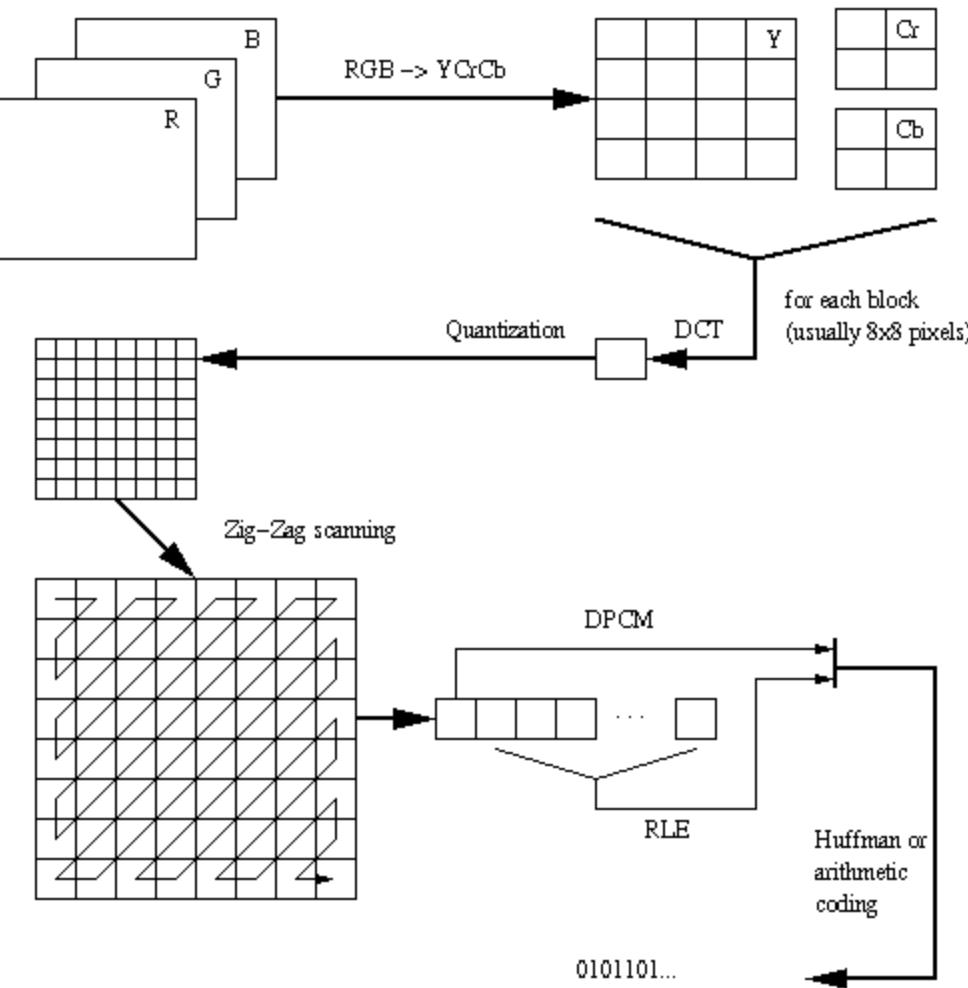
$$Y = 0.299 R + 0.587G + 0.114 B$$

$$Cb = -0.1687R - 0.3313G + 0.5B + 128$$

$$Cr = 0.5 R - 0.4187G - 0.0813 B + 128$$



# 彩色图像 JPEG 压缩流程

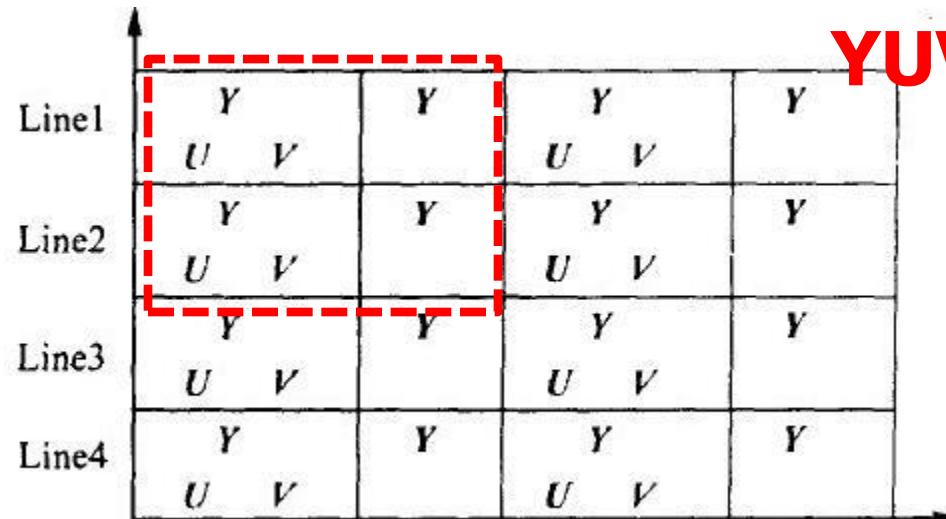


RGB → YUV 时，空间采样对 Y、U、V 三个分量采用不同的采样频率。

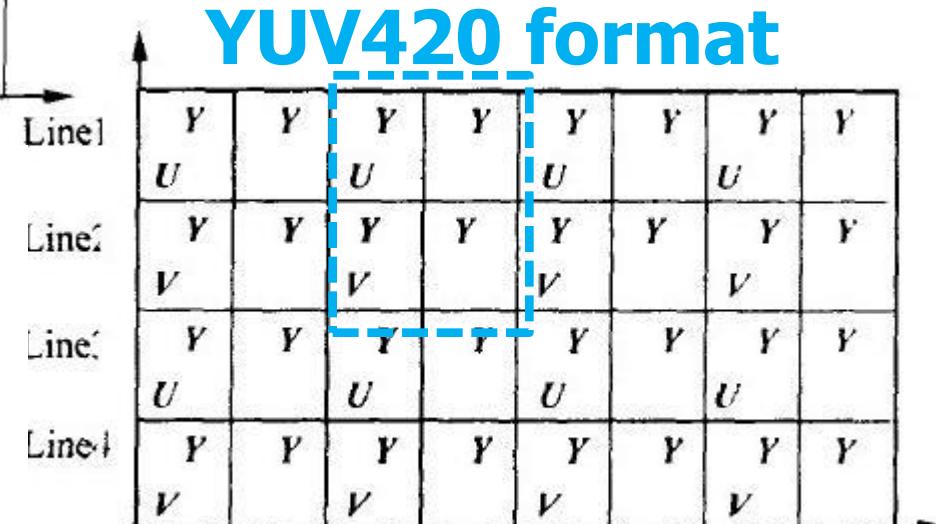
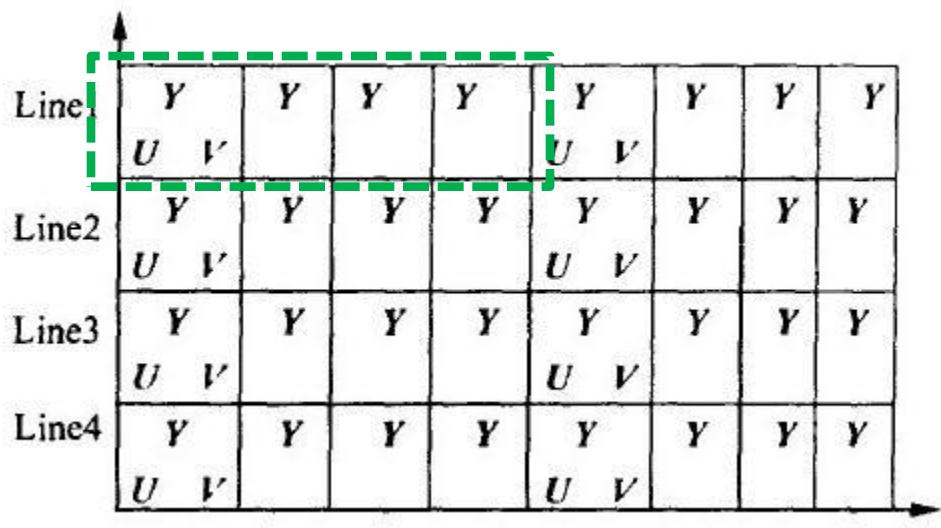
YUV 4:2:2  
YUV 4:2:0  
YUV 4:1:1



# 彩色图像的采样格式



**YUV422 format**



**YUV420 format**

{yhh,cxh}@ustc.edu.cn



# JPEG小结

◆JPEG标准中以DCT为基础的编码过程，实际上是在消除图象中的相关性，或者说消除图象中的冗余度，这种冗余度包括下列三种。

- 第一种是编码冗余度。例如，DCT变换，哈夫曼编码，是消除编码冗余度。
- 第二种是象素间冗余度。例如，直流系数用差分编码就是消除相邻子图象间的灰度(或亮度)冗余度。
- 第三种是心理视觉冗余度。例如，用量化过程，就是利用人眼对各种空间频率，包括亮度、色度、纵、横方向的高频、低频的敏感程度不同，从而降低和消除一部分数据，达到数据压缩的目的，或降低传输位率，同时又不损害心理视觉对图象的主观评价。这就是充分利用心理视觉冗余度对图象数据进行压缩。



# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
  - 图像和视频的冗余
  - JPEG标准
    - JPEG标准的三个基本要素
    - JPEG中的有损和无损压缩
    - JPEG标准的工作模式
  - JPEG标准中的压缩算法
    - JPEG算法概要
    - 彩色图像 JPEG压缩
  - **JPEG 2000/JPEG XR**
  - H.264 Intra Frame
- ◆ 2.5 视频压缩的国际标准
- ◆ 2.6 可伸缩性编码和分布式编码



# JPEG-2000标准：设计目标

◆ 随着多媒体应用领域的快速增长，传统JPEG压缩技术已无法满足人们对数字化多媒体图像资料的要求。针对这些问题，专家们开始了下一代JPEG 2000标准的制定，最终标准于2000年12月出台。JPEG 2000的预期目标：

高压缩率 有损/无损压缩  
渐进传输 感兴趣区域压缩  
色彩模式 图像处理简单



# JPEG-2000标准：规范内容

◆JPEG2000主要由6部分组成。其中第一部分是编码的核心部分，具有相对而言最小的复杂性，可以满足80%的应用需求，其地位相当于JPEG的基本系统，也是公开并可免费使用的。

- 规定了解码过程，将压缩图像数据转换为重建图像数据
- 规定了码流语法，其中包含对压缩图像数据的解释信息
- 规定了JP2文件格式
- 提供了编码过程指导，可将原图像转换为压缩图像数据
- 提供了实际进行编码处理的实现的指导

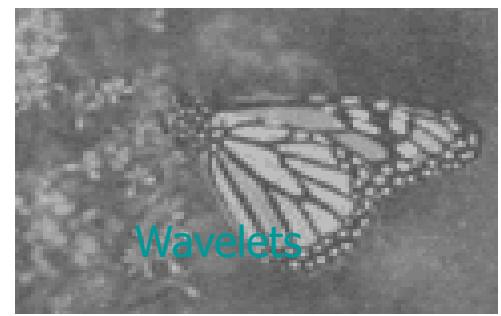
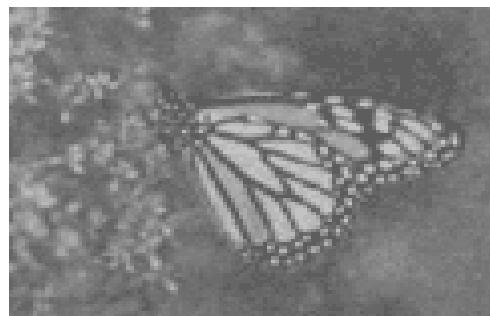
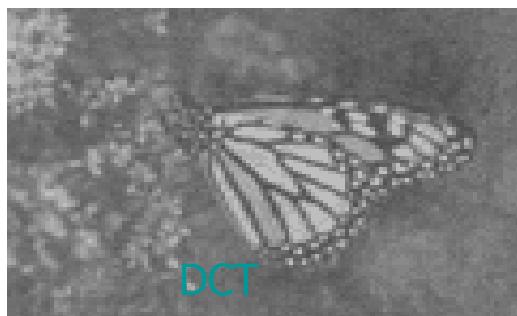
◆2~6部分定义了压缩文件格式的扩展部分,以满足特殊应用，其复杂度大大增加。其中包括：

- 编码扩展(第二部分)；
- Motion JPEG2000（MJP2，第三部分）；
- 一致性测试（第四部分）；
- 参考软件（第五部分）；
- 混合图像文件格式（第六部分）。



# JPEG-2000标准：技术特点

- ◆ JPEG2000是基于小波变换的图像压缩标准。扩展名“.jp2”
- ◆ 在高压缩比的情形下，JPEG2000图像失真程度一般会比传统的JPEG图像要小。JPEG2000同时支持有损压缩和无损压缩。另外，JPEG2000也支持更复杂的渐进式显示和下载。
- ◆ 大多数的浏览器仍然没有缺省支持JPEG2000图像文件的显示。但是，由于JPEG2000在无损压缩下仍然能有比较好的压缩率，所以JPEG2000在图像质量要求比较高的医学图像的分析和处理中已经有了一定程度的广泛应用。





# JPEG-2000标准： 离散小波变换（DWT）

□ 小波变换实现多分辨率的分等级显示和渐进显示：

$$I_a = 2LL;$$

$$I_b = 2LL + 2HL + 2LH + 2HH;$$

$$I_c = 2LL + 2HL + 2LH + 2HH \\ + 1HL + 1LH + 1HH;$$

$$I_d = I_a (\uparrow 4)$$

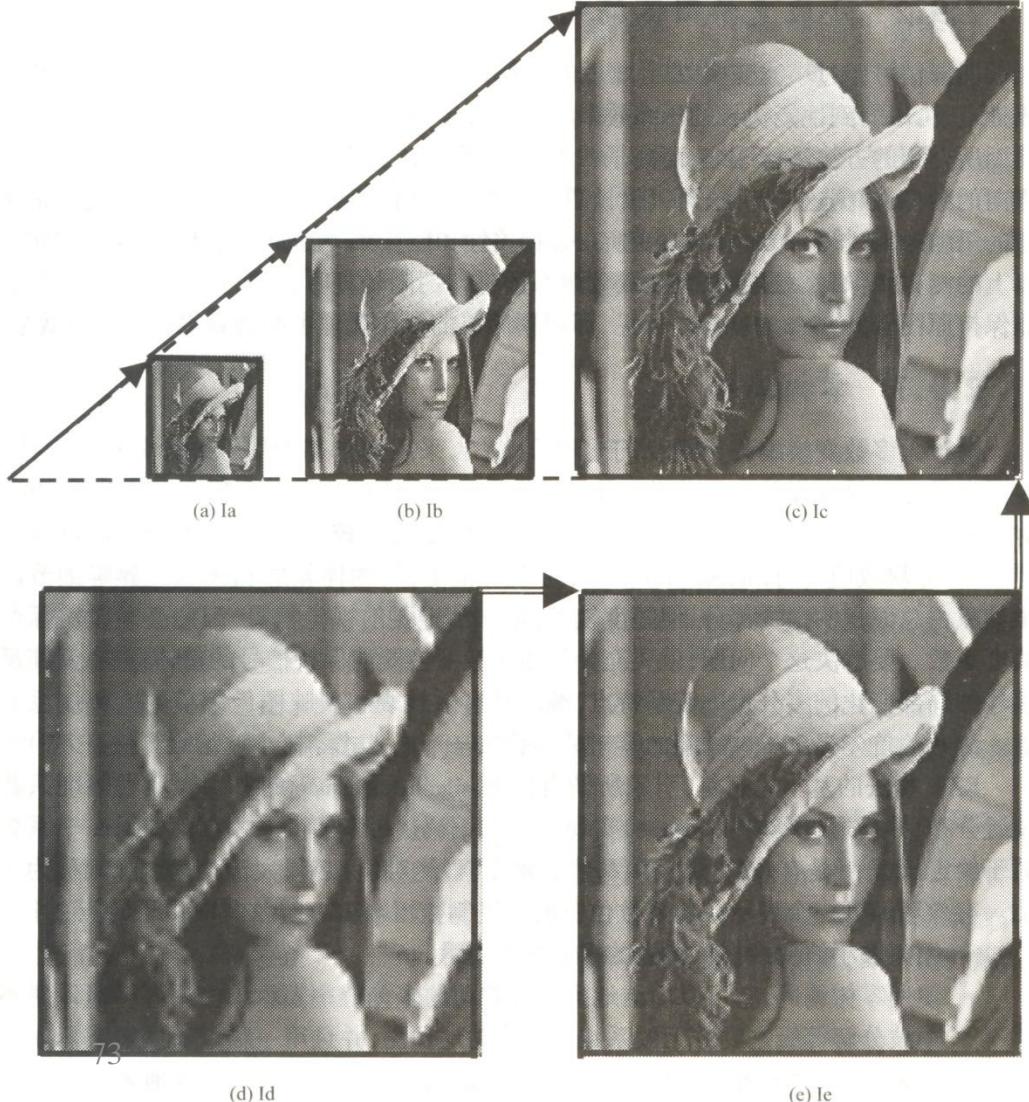
对  $I_a$  水平垂直各上抽样 1: 4

$$I_e = I_b (\uparrow 2)$$

对  $I_b$  水平垂直各上抽样 1: 2

分等级显示：  $I_a, I_b, I_c$ ；

渐进显示：  $I_d, I_e, I_c$ ；





# JPEG-2000标准： 离散小波变换（DWT）

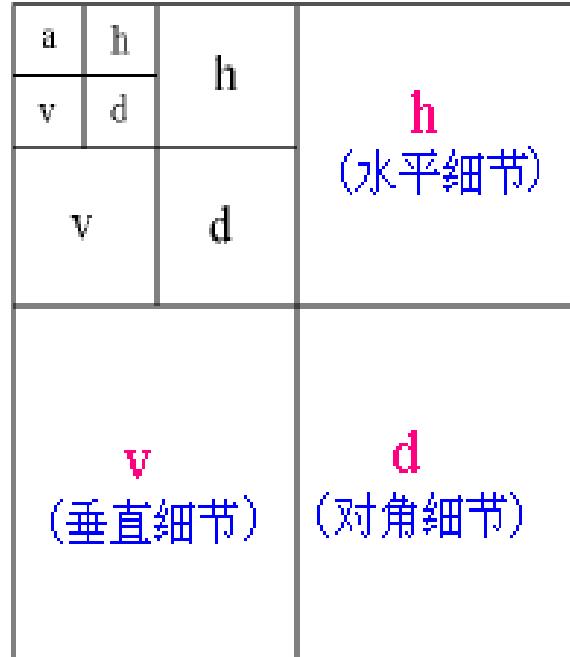
**a**  
(近似值)

**h**  
(水平细节)

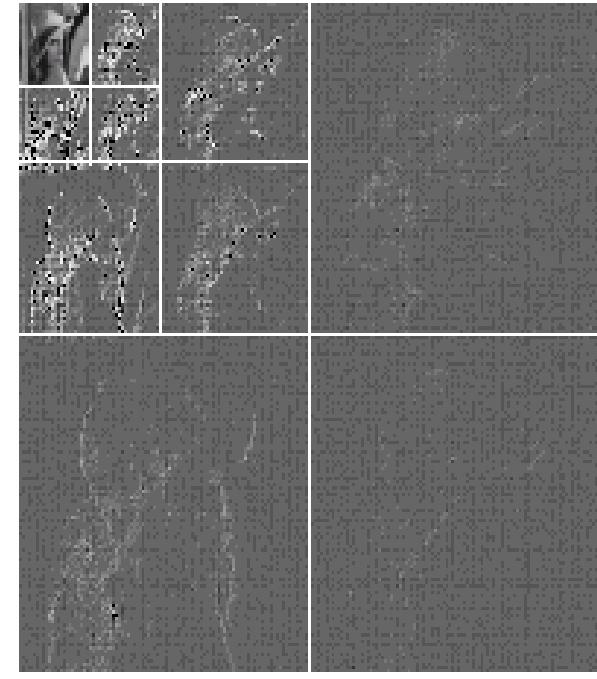
**v**  
(垂直细节)

**d**  
(对角细节)

(a) 一级分解



(b) 三级分解

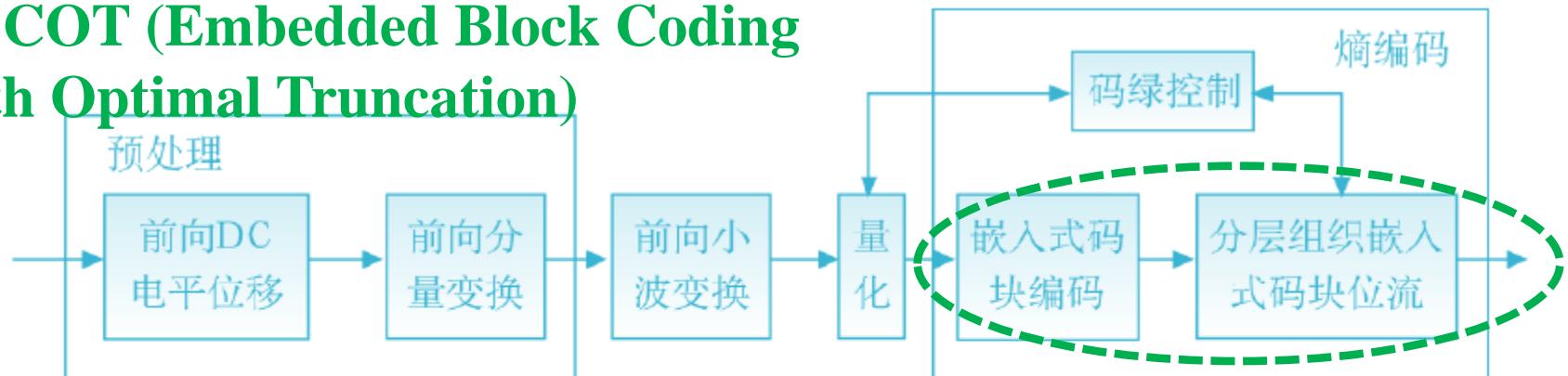




# JPEG2000编解码器框架

熵编码：首先使用**基于上下文的算术编码器**，每个码块进行独立的嵌入式码块编码；然后将所有码块的压缩位流适当截取，组织成具有**不同质量级**的压缩位流层。每一层上的压缩位流连同其前面的所有层的压缩位流，可重构出一定质量的图像。

## EBCOT (Embedded Block Coding with Optimal Truncation)





# JPEG-2000标准： JPEG2000和JPEG的效果比较

- ◆ 在有损压缩下，JPEG2000一个比较明显的优点就是没有JPEG压缩中的**马赛克失真**效果。JPEG2000的失真主要是模糊失真。**模糊失真**产生的主要原因是在编码过程中高频量一定程度的衰减。传统的JPEG压缩也存在模糊失真的问题。
- ◆ 在低压缩比情形下（比如压缩比小于10: 1），传统的JPEG图像质量有可能要比JPEG2000要好。JPEG2000在压缩比比较高的情形下，优势才开始明显。整体来说，和传统的JPEG相比，JPEG2000仍然有很大的技术优势，通常压缩性能大概可以提高20%以上。一般在压缩比达到100: 1的情形下，采用JPEG压缩的图像已经严重失真并开始难以识别了，但JPEG2000的图像仍可识别。

JPEG2000存在版权和专利的风险。这是目前JPEG2000技术没有得到更广泛应用的原因之一



# JPEG-2000标准： JPEG2000和JPEG的效果示例1



Barbara.bmp (50% size)

测试图像 Barbara

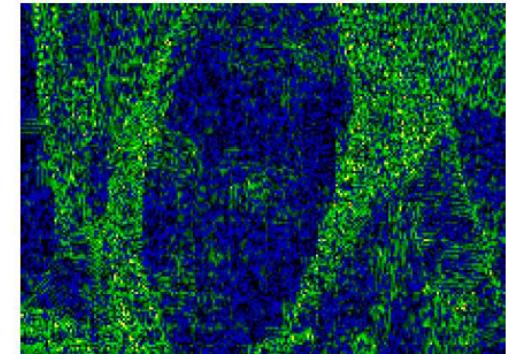
Black and white image. Main feature – stripes on tablecloth, on scarf and on pants where moire and other artifacts often appear.



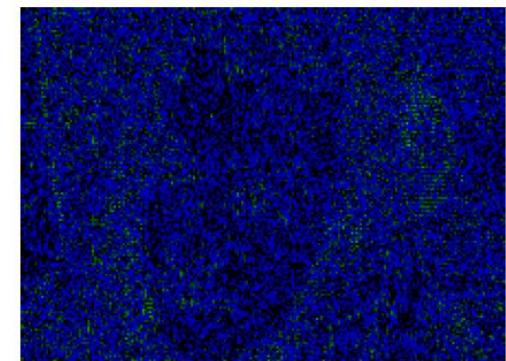
Barbara, JPEG, 31561 bytes



Barbara, JPEG 2000, ACDSee, 32412 bytes



Difference between JPEG and original



Difference between JPEG 2000 and original



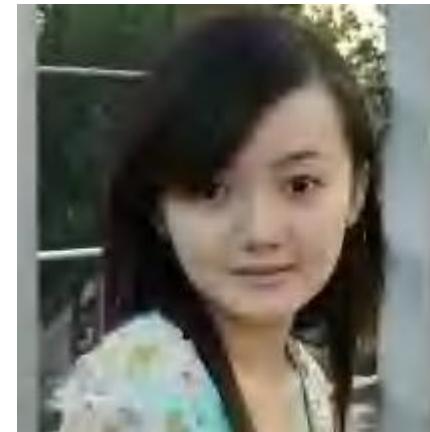
# JPEG-2000标准： JPEG2000和JPEG的效果示例2



原图 (48.9KB)



JPEG压缩 (1.85KB) JPEG2000压缩(1.79KB)





# JPEG-2000标准： JPEG2000和JPEG的效果示例3



压缩率再高，  
也只有模糊  
失真





# JPEG-2000标准： JPEG2000和JPEG的效果示例4

◆ 基于感兴趣区域的压缩，左边为原图，右边为JPEG局部压缩后的图像





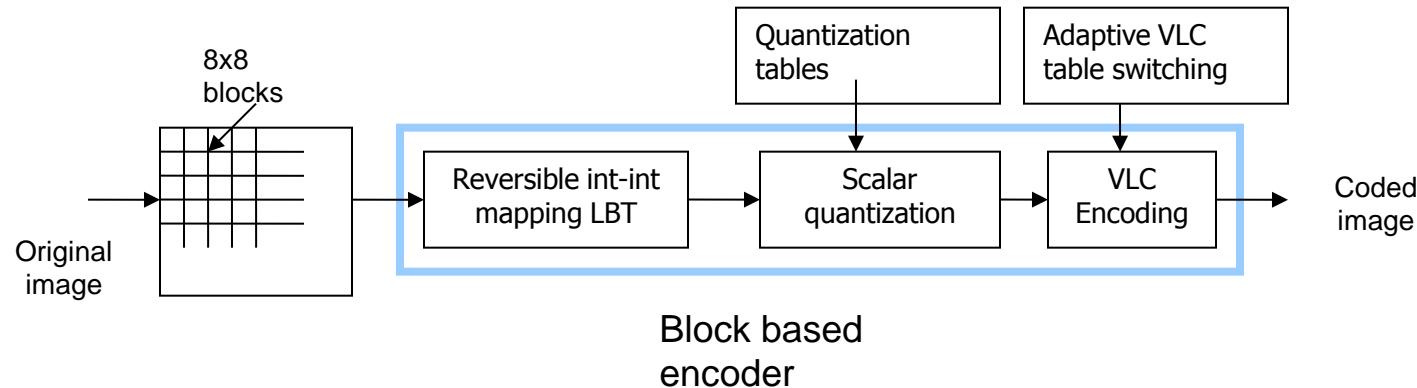
# JPEG-XR(JPEG extended range)

◆JPEG XR（前身为Windows媒体照片和高清照片，**HD Photo**）是一个静止图像压缩标准和文件格式的照片图像连续色调的基础上，开发和技术最初由微软的专利作为Windows媒体系列的一部分。Windows Vista版本以上，已经内置支持JPEG XR格式。扩展名为.wdp的图片，会在资源管理器中自然的得到预览。Internet Explorer 9也支持对.wdp图片进行预览。

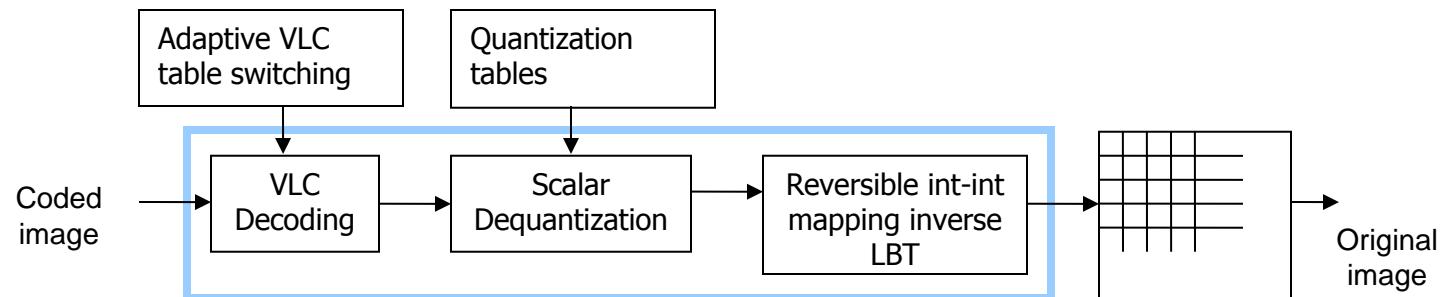




# JPEG-XR encoder and decoder



- **lapped bi-orthogonal transform (LBT)**
- **VLC Encoding**

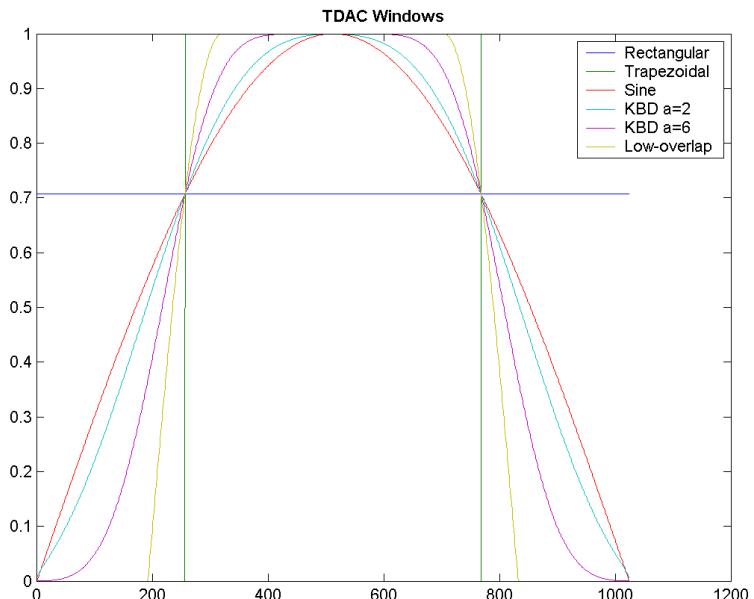
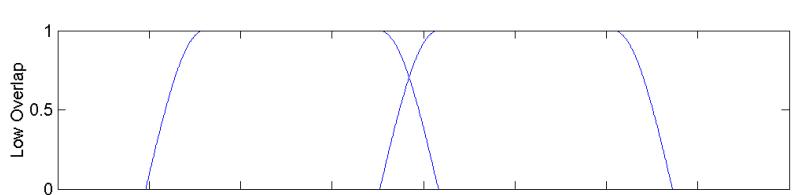
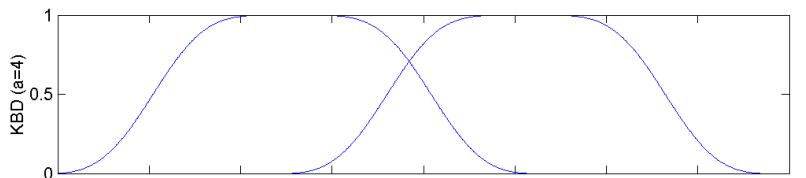
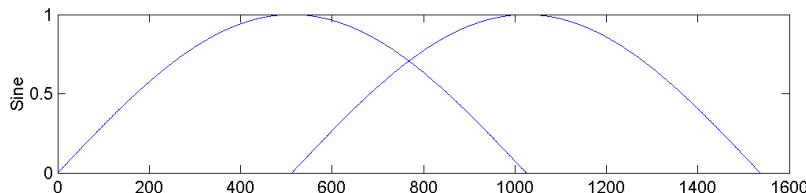




# Kaiser-Bessel Derived (KBD) window

## Review

- ◆ A 50% overlap add (OLA) structure with certain pre and post, time domain aliasing cancellation (TDAC) windowing, the initial signal can be completely recovered.

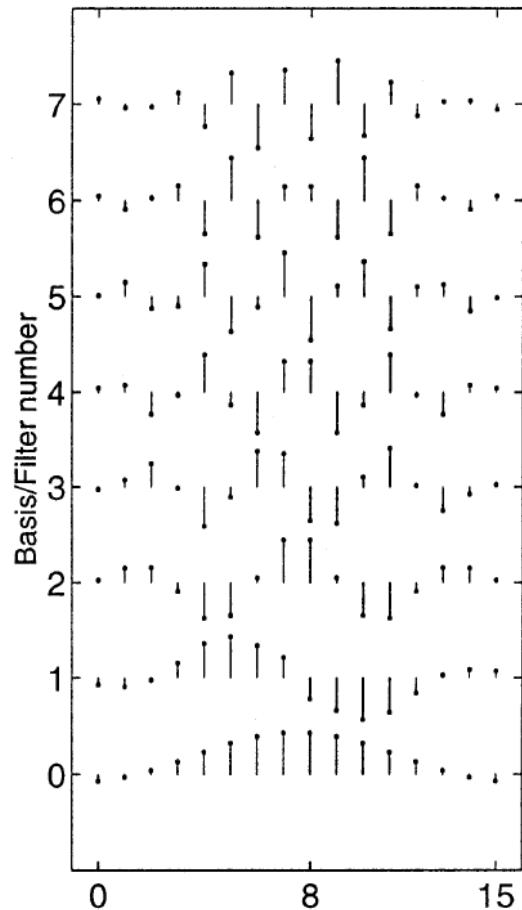
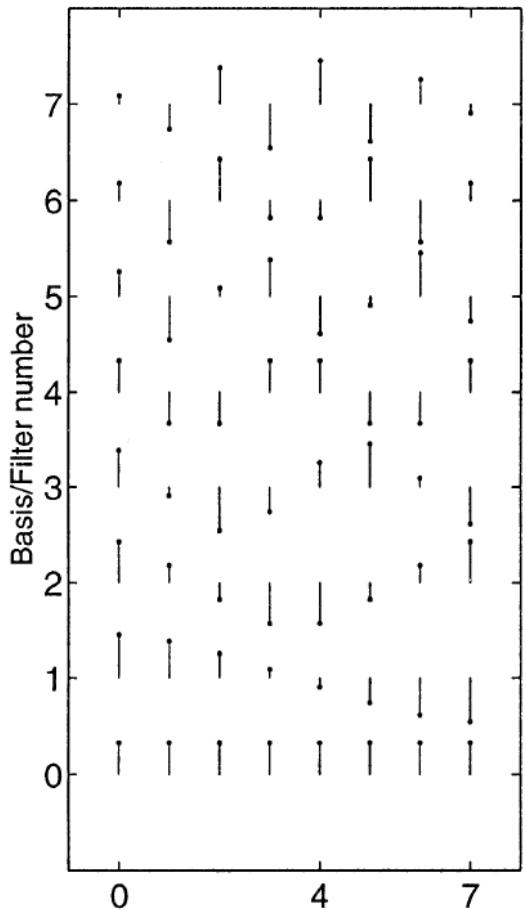




# 正交交叠变换

## LOT: Orthogonal Lapped Transforms

Bases for the 8-point DCT ( $M = 8$ ) (left) and for the LOT (right) with  $M = 8$ .





# 性能评价

Subjective evaluation of JPEG XR image compression  
De Simone, Francesca; Goldmann, Lutz; Baroncini, Vittorio; Ebrahimi, Touradj  
Published in: Proceedings of SPIE 2009, vol. 7443

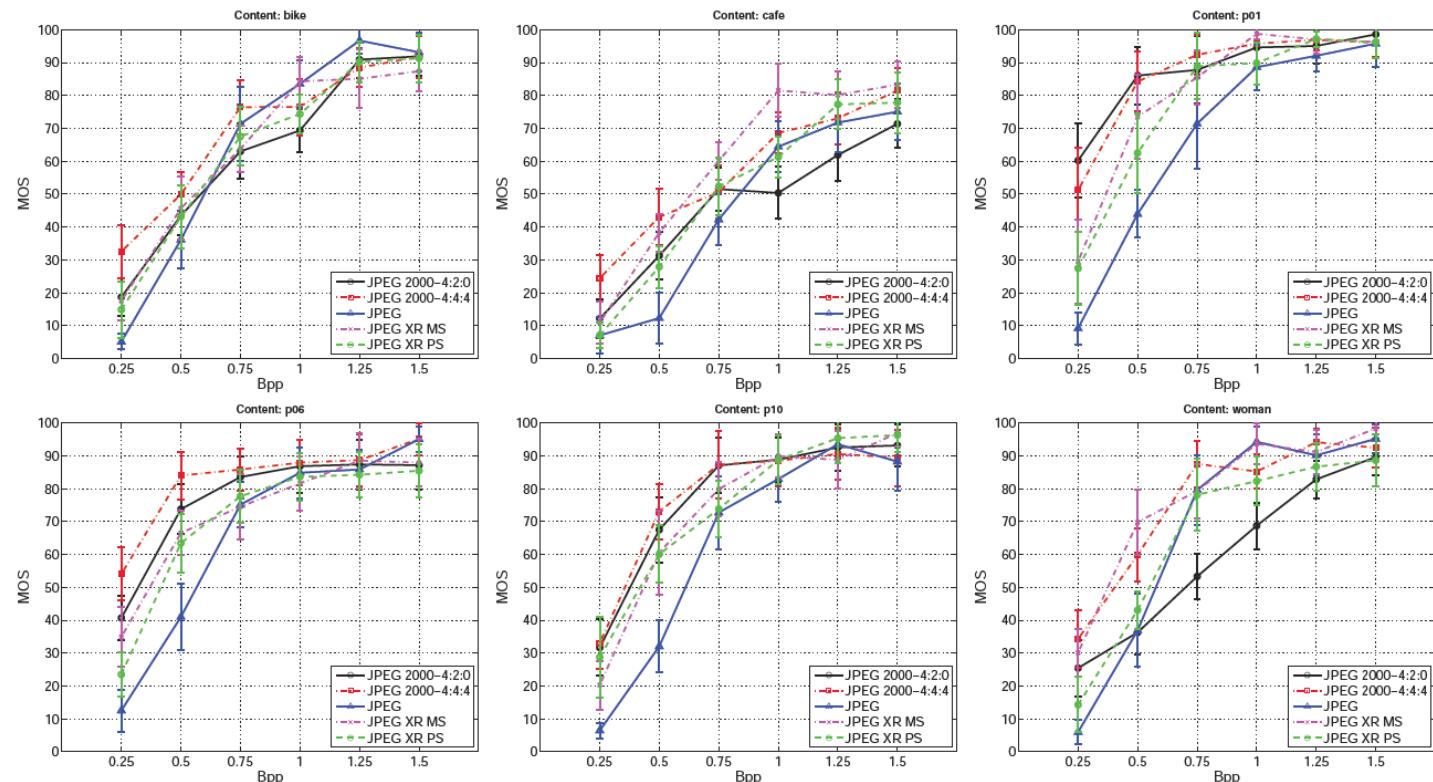


Figure 5: Mean opinion score vs. bitrate results for the different codecs across test images. From top left to bottom right: *bike*, *cafe*, *p01*, *p06*, *p10* and *woman*.



(a) p04

(b) p14

(c) p22

(d) p30

(e) p01

(f) p06

(g) p10

(h) bike

(i) cafe

(j) woman



# 小结：JPEG2000/JPEG XR

## ◆ 变换

- DCT → DWT → LBT

## ◆ 熵编码

- Huffman编码 → VLC

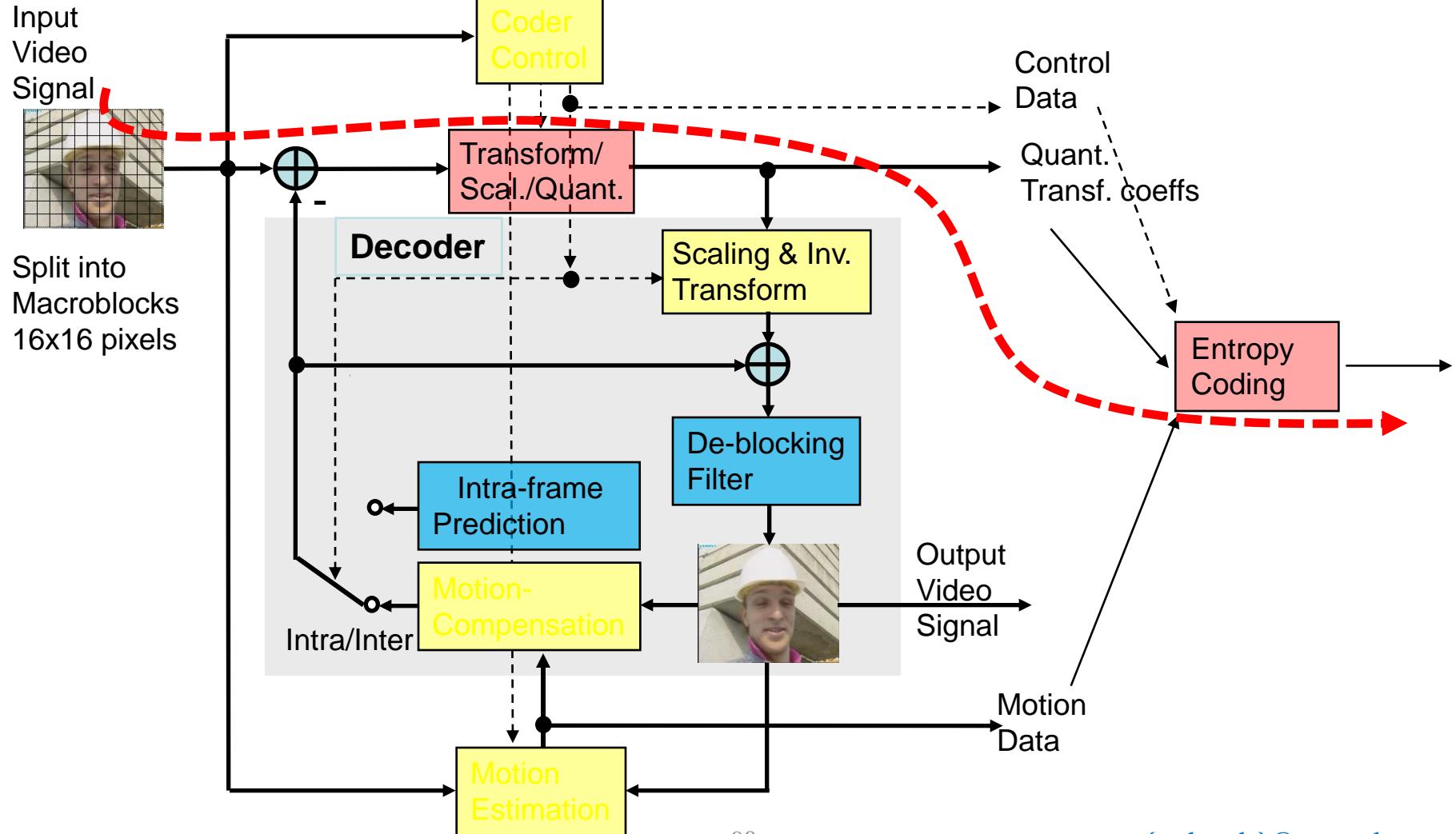


# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
  - 图像和视频的冗余
  - JPEG标准
    - JPEG标准的三个基本要素
    - JPEG中的有损和无损压缩
    - JPEG标准的工作模式
  - JPEG标准中的压缩算法
    - JPEG算法概要
    - 彩色图像 JPEG压缩
  - JPEG 2000/JPEG XR
  - H.264 Intra Frame
- ◆ 2.5 视频压缩的国际标准
- ◆ 2.6 可伸缩性编码和分布式编码



# H.264对静图像的编码

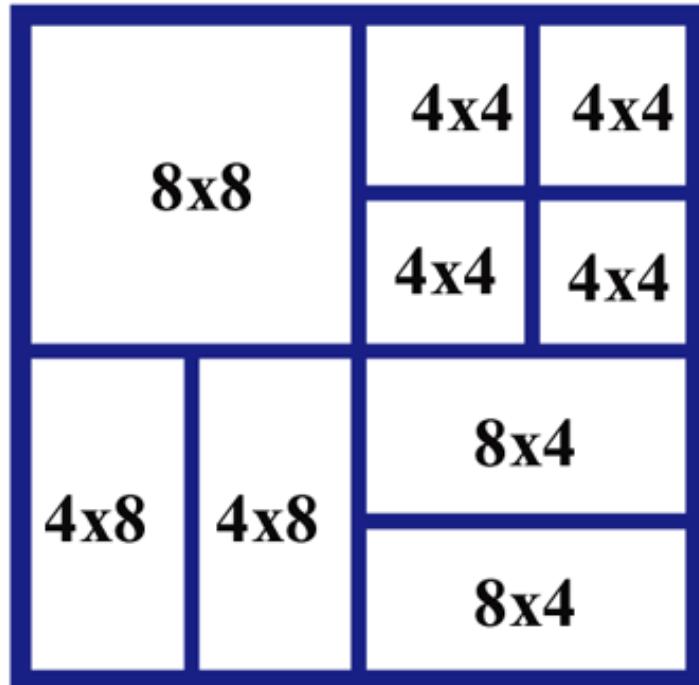




# H.264对静图像编码时的宏块

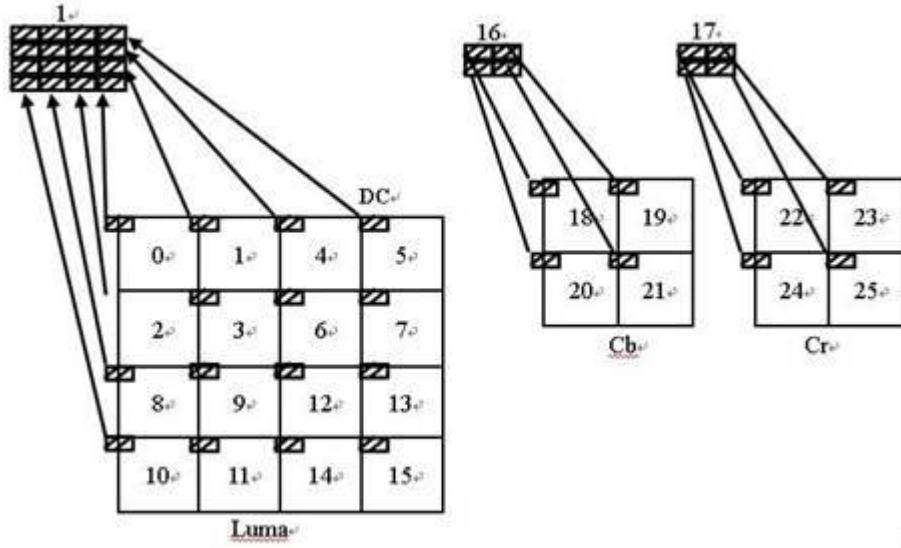
2003

- ◆ 在JPEG中，每个宏块的大小固定为 $8*8$
- ◆ H.264中宏块大小为 $16*16$ 
  - 可以对不同大小和形状的宏块进行运动矢量的估计



# H.264对静图像编码时的变换

## ◆ 4x4 Block Integer Transform



$$H = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

宏块大小为  $16 \times 16$ ，对其中每个  $4 \times 4$  大小的块进行上述  $4 \times 4$  的 DCT 变换后，得到 16 个  $4 \times 4$  的变换矩阵。为了进一步提高压缩效率，还允许把每个  $4 \times 4$  的变换矩阵中的直流分量 DC，单独取出组成一个新的  $4 \times 4$  矩阵，对此矩阵进行 Hardamard 变换。



# H.264对静图像编码时的熵编码

- ◆ H.264建议同时采用了两种熵编码模式
- ◆ 基于上下文的二进制算术编码CABAC
  - CABAC: Context-Adaptive Binary Arithmetic Coding
- ◆ 可变长编码VLC: CAVLC
  - CAVLC: Context-Adaptive Variable Length Coding

# H.264对静图像编码的去块效应滤波

◆去块效应滤波的作用就是用来消除解码图像中的块效应。块效应产生的原因是各个宏块分别进行量化，这样在相邻宏块的交界处，因量化步长不同而导致原本很接近的像素值重构后产生了较大的差异，形成明显的块边界。去块效应滤波是在 $4\times 4$ 的块边界上滤波，使块边界趋于平滑。





# H.264对静图像编码的帧内预测

- ◆ Intra-frame Prediction
- ◆ JPEG中每个宏块都进行编码
- ◆ 相邻宏块之间通常具有很强的相关性（空间冗余）
- ◆ 后编码的宏块若和已编码过的宏块相似，则只需要记录预测值和实际值之间的差值

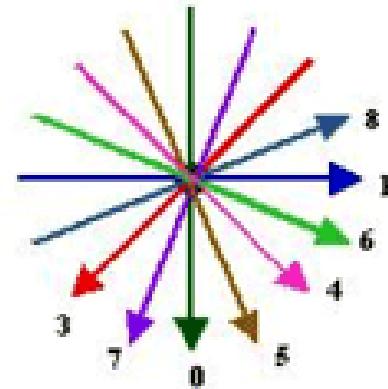
| Samples that are already encoded |   | Samples to be intra predicted |   |   |   |   |   |   |
|----------------------------------|---|-------------------------------|---|---|---|---|---|---|
| M                                | A | B                             | C | D | E | F | G | H |
| I                                | a | b                             | c | d |   |   |   |   |
| J                                | e | f                             | g | h |   |   |   |   |
| K                                | i | j                             | k | l |   |   |   |   |
| L                                | m | n                             | o | p |   |   |   |   |



# 帧内预测的模式

- ◆ 对宏块的编码从左上角开始
- ◆ 包括9种 $4 \times 4$ 亮度块的预测、4种 $16 \times 16$ 亮度块的预测和4种色度块的预测

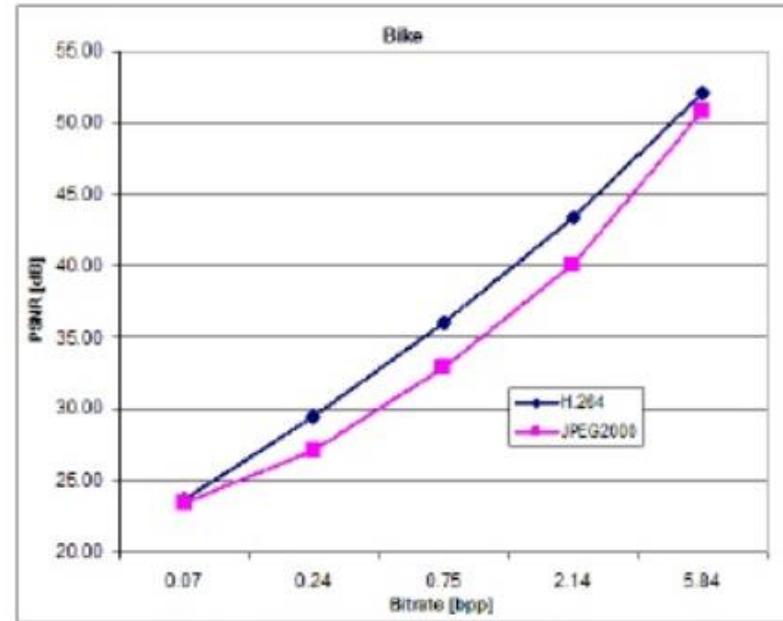
| M | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| I | a | b | c | d |   |   |   |   |
| J | e | f | g | h |   |   |   |   |
| K | i | j | k | l |   |   |   |   |
| L | m | n | o | p |   |   |   |   |



Samples a, b, ..., p are predicted from samples A, ..., M that have been encoded previously.



# H.264静图像对比JPEG2000: PSNR



Giacomo Camperi, Vittorio Picco

H.264/AVC intra coding and JPEG 2000 comparison, April 11, 2008

<http://zh.scribd.com/doc/2581731/H264AVC-intra-coding-and-JPEG-2000-comparison>

Figure 1: “Bike” test image

| QP | Bitrate [bpp] | PSNR H.264 [dB] | PSNR JPEG [dB] | Compression ratio |
|----|---------------|-----------------|----------------|-------------------|
| 10 | 5.84          | 52.11           | 50.85          | 4:1               |
| 20 | 2.14          | 43.40           | 40.14          | 11:1              |

Table 1: High bitrate tests for the image Bike

| QP | Bitrate [bpp] | PSNR H.264 [dB] | PSNR JPEG [dB] | Compression ratio |
|----|---------------|-----------------|----------------|-------------------|
| 30 | 0.75          | 36.08           | 32.91          | 32:1              |
| 40 | 0.24          | 29.47           | 27.14          | 100:1             |

Table 2: Medium bitrate tests for the image Bike

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{255^2}{\sum_{i=0}^N (X_i - Y_i)^2} \right) [\text{dB}]$$

{ynh,cxh}@ustc.edu.cn



# H.264静图像对比JPEG2000: PSNR



Figure 7: Particular of “Space” QP=50: original (center), JPEG 2000 (left), H.264 (right)

| QP | Bitrate [bpp] | PSNR H.264 [dB] | PSNR JPEG [dB] | Compression ratio |
|----|---------------|-----------------|----------------|-------------------|
| 40 | 0.04          | 38.39           | 38.62          | 578:1             |
| 50 | 0.02          | 30.46           | 35.08          | 1120:1            |

Giacò

H.264

<http://zh.scribd.com/doc/2581731/H264AVC-intra-coding-and-JPEG-2000-comparison>



Table 4: Low bitrate tests for the image Space

## Space

|           |                 |
|-----------|-----------------|
| Width     | 1024            |
| Height    | 512             |
| File size | 1,572,864 bytes |

并非所有情况下  
H.264质量都好

{ynh,cxh}@ustc.edu.cn



# 在H.264十年之后

International Telecommunication Union

ITU-T

TELECOMMUNICATION  
STANDARDIZATION SECTOR  
OF ITU

H.265  
(04/2013)

Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC)  
IEEE Transactions on Circuit and System for Video Technology, vol. 22, N0. 12, Dec 2012

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS  
Infrastructure of audiovisual services – Coding of moving video

High efficiency video coding

Recommendation ITU-T H.265



HEVC Complexity and Implementation Analysis

IEEE Transactions on Circuit and System for Video Technology, vol. 22, N0. 12, Dec 2012

**HEVC encoders can achieve equivalent subjective reproduction quality as encoders that conform to H.264/MPEG-4 AVC when using approximately 50% less bit rate on average.**

**Overall the complexity of HEVC decoders does not appear to be significantly different from that of H.264/AVC decoders. HEVC encoders are expected to be several times more complex than H.264/AVC encoders, and a subject of research in years to come.**

{ynh,cxh}@ustc.edu.cn



# HEVC(H.265)对静图像的编码

2013.04

- ◆ 宏块大小:  $L \times L$ ,  $L = 16, 32$ , or  $64$  samples
- ◆ **Integer** basis functions similar to those of a discrete cosine transform (DCT) are defined for the square transform blocks sizes  $4 \times 4$ ,  $8 \times 8$ ,  $16 \times 16$ , and  $32 \times 32$ .
- ◆ **Intrapicture prediction**: supports **33 directional modes**
- ◆ **Entropy coding**: Context adaptive binary arithmetic coding (CABAC) is used for entropy coding. This is similar to the CABAC scheme in H.264/MPEG-4 AVC, but has undergone several improvements to improve its throughput speed (especially for **parallel-processing architectures**) and its compression performance, and to reduce its context memory requirements.

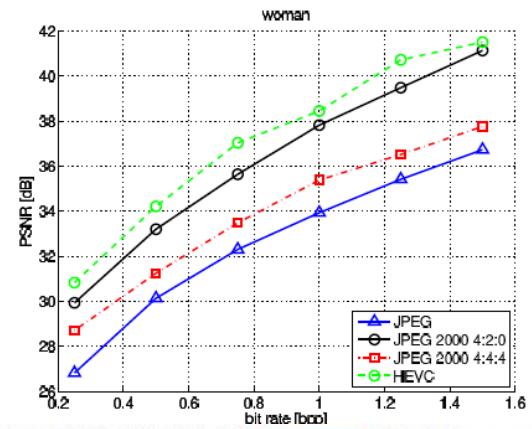
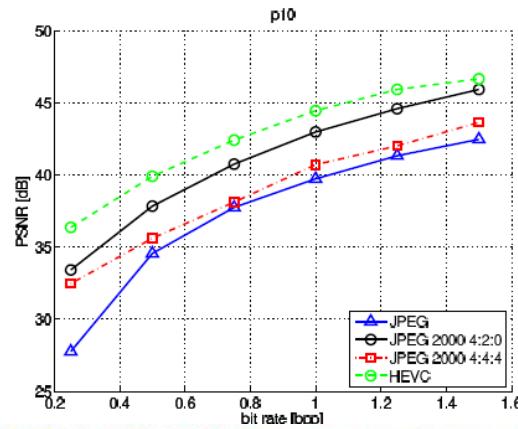
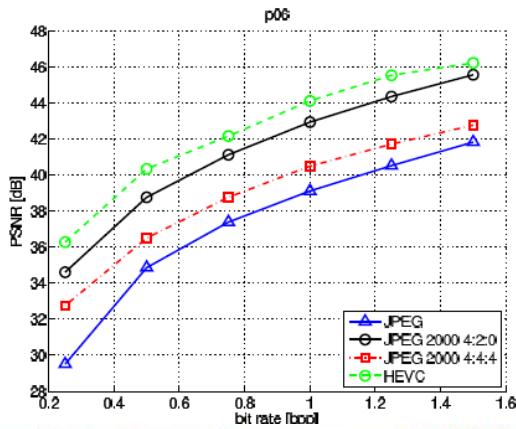
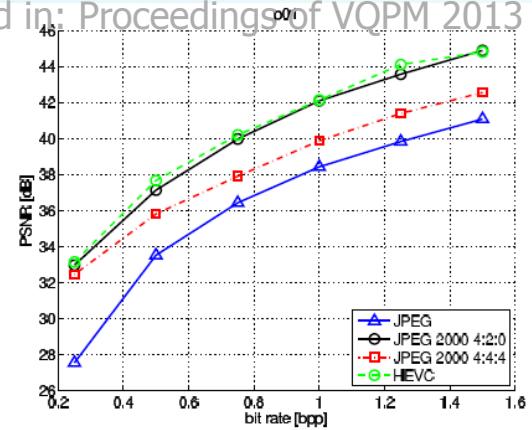
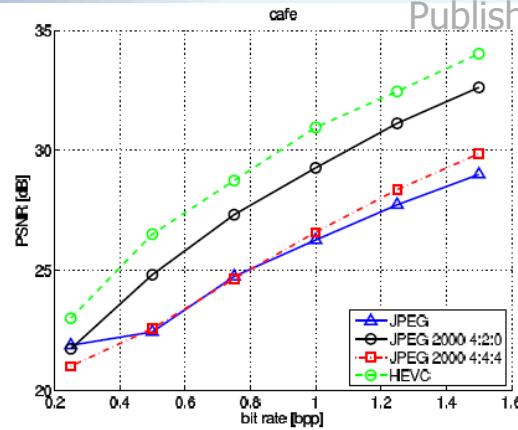
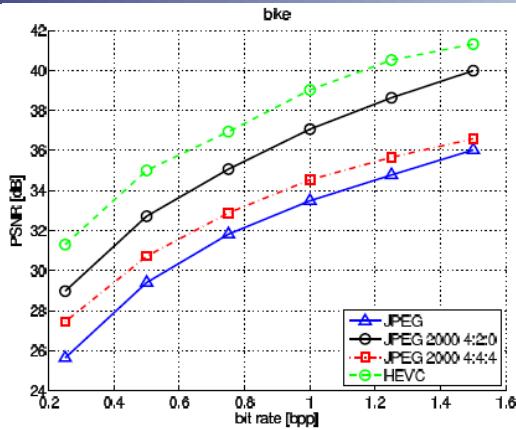


# PSNR: HEVC/JPEG/JPEG2000

Subjective Evaluation of HEVC Intra Coding for Still Image Compression

Philippe Hanhart, etc.

Published in: Proceedings of VQPM 2013



(a) p04

(b) p14

(c) p22

(d) p30

(e) p01

(f) p06

(g) p10

(h) bike

(i) cafe

(j) woman

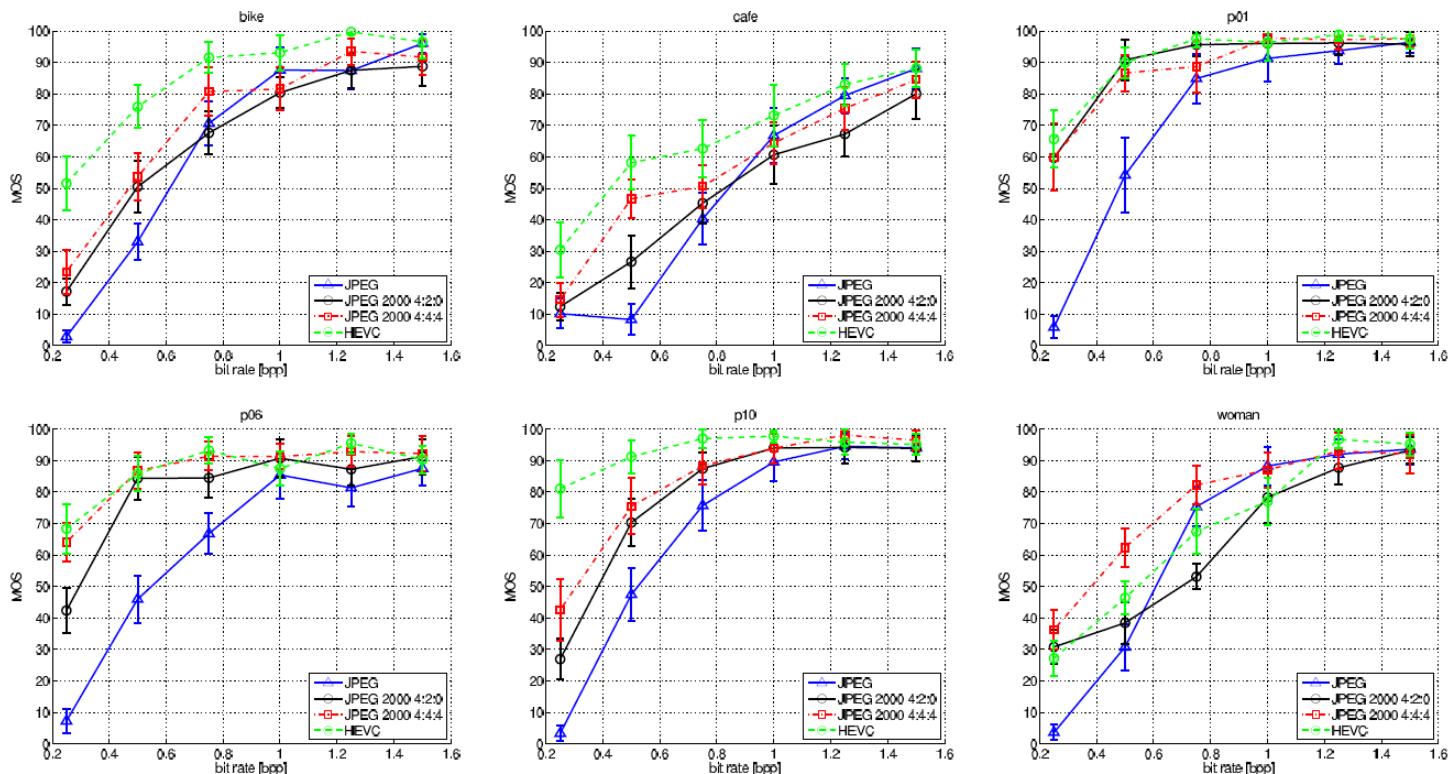


# MOS: HEVC/JPEG/JPEG2000

Subjective Evaluation of HEVC Intra Coding for Still Image Compression

Philippe Hanhart, etc.

Published in: Proceedings of VQPM 2013



(a) p04

(b) p14

(c) p22

(d) p30

(e) p01

(f) p06

(g) p10

(h) bike

(i) cafe

(j) woman



# 小结：H.264/HEVC图像压缩的新意

## ◆ 变换

- DCT(1992) → DWT(2000) → LBT(2009)
- 视频编码标准中的图像编码 → Integer Transform(2003)

## ◆ 熵编码

- Huffman编码(1992) → VLC(2009)
- 视频编码标准中的图像编码 → Context-Adaptive VLC(2003)

## ◆ H.264 (2003) 帧内预测

## ◆ H.264 (2003) 去块效应滤波器

## ◆ HEVC (2013) 考虑算法的实现结构

- parallel-processing architectures



# 小结：20年来图像压缩相关的标准

## ◆ 图像

- ◆ JPEG, 1992
- ◆ JPEG-LS(1998)
- ◆ JPEG 2000, 2000
- ◆ JPEG XR, 2009

## ◆ 视频

- ◆ H.261, 1988
- ◆ MPEG-1, 1993
- ◆ MPEG-2/H.262, 1996
- ◆ H.263, 1996
- ◆ MPEG-4, 1999
- ◆ **MPEG-4 AVC/H.264, 2003**
- ◆ HEVC, 2013年4月13号

## ◆ 变换

- DCT、DWT、LBT
- Integer Transform

## ◆ 熵编码

- Huffman、VLC
- Context-Adaptive VLC

## ◆ H.264

- 去块效应滤波器
- 帧内预测

## ◆ HEVC(H.265)

- 考虑算法的实现结构



下次上课时间：10月11日