

本节为[opencv数字图像处理](#) (14) : **图像压缩**的第一小节, **图像压缩基础知识**, 主要包括: 三种数据冗余、图像信息的度量、压缩的保真度、压缩/解压模型以及常见的压缩标准、格式和容器。

数据压缩比是指减少表示给定信息量所需数据量的处理, 系统数量的信息可以用不同数量的数据表示, 包含不相关或重复信息的表示称之为冗余数据。令 $b$ 和 $b'$ 代表相同信息的两种表示中的比特数, 那么 $b$ 比特表示的相对数据冗余 $R$ 为 $R = 1 - 1/C$ , 其中 $C$ 为压缩率, 定义为 $C = b/b'$ 。若 $C = 10$ , 那么相对数据冗余为0.9, 表明90%的数据是冗余的。

二维灰度阵列受三种主要类型的数据冗余影响:

- 编码冗余
- 空间和时间冗余
- 不相关的信息: 被人类视觉系统忽略或者与用途无关的信息 当一个或多个冗余被减少或消除时, 压缩就实现了。

## 1. 编码冗余

假设 $[0, L - 1]$ 内的一个离散随机变量 $r_k$ 用来表示一幅 $M \times N$ 的图像的灰度, 并且每个 $r_k$ 发生的概率为 $p_r(r_k)$ , 则 $p_r(r_k) = \frac{n_k}{MN}$ , 其中 $L$ 是灰度级数,  $n_k$ 是第 $k$ 级灰度在图像中出现的次数, 若每个 $r_k$ 值的比特数为 $l(r_k)$ , 则每个像素所需的平均比特数为 $L_{avg} = \sum_{k=0}^{L-1} l(r_k) p(r_k)$ , 表示大小为 $M \times N$ 的图像所需的总比特数为 $MNL_{avg}$ , 如果用 $m$ 比特的固定长度的码表示灰度, 即 $l(r_k) = m$ 时,  $L_{avg} = m$ , 当然也可以用 $l_2(r_k)$ 的变长编码, 如下所示:

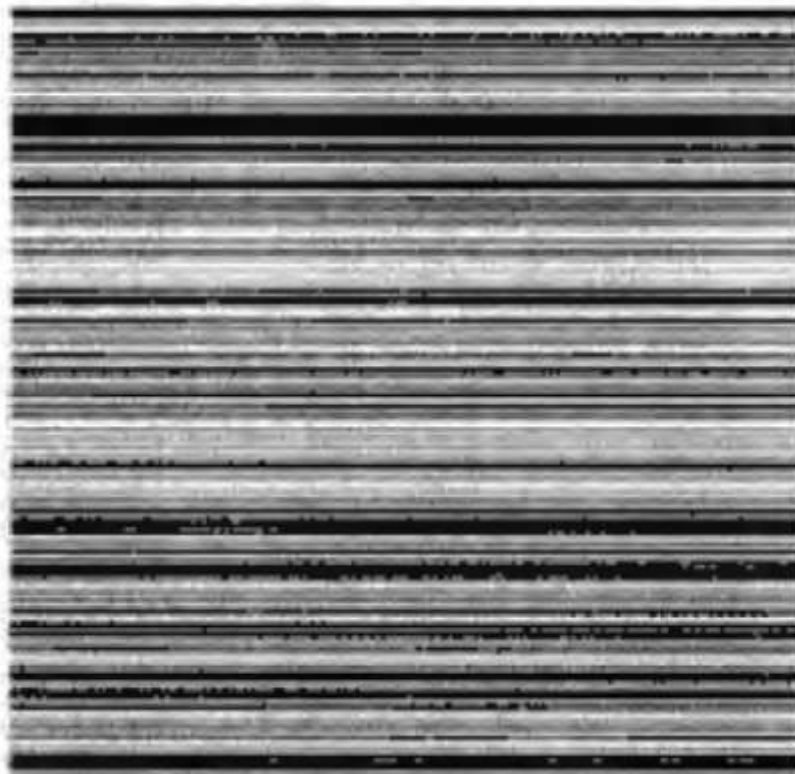
$r_k$	$p_r(r_k)$	编码 1	$l_1(r_k)$	编码 2	$l_2(r_k)$
$r_{87} = 87$	0.25	01010111	8	01	2
$r_{128} = 128$	0.47	10000000	8	1	1
$r_{186} = 186$	0.25	11000100	8	000	3
$r_{255} = 255$	0.03	11111111	8	001	3
$r_k, k \neq 87, 128, 186, 255$	0	—	8	—	0

则 $L_{avg} = 0.25(2) + 0.47(1) + 0.25(3) + 0.03(3) = 1.81$  比特, 整幅图像  
 $MNL_{avg} = 256 \times 256 \times 1.81 = 1186210$ , 相比于定长编码, 压缩率和相对冗余为  
 $C = \frac{256 \times 256 \times 8}{1186210} = 4.42$  和  $R = 1 - \frac{1}{4.42} = 0.774$ 。

如上所示, 不去全部时间概率的优势, 就会出现编码冗余, 当用自然二进制码表示灰度时(固定码长), 编码冗余总数存在的。

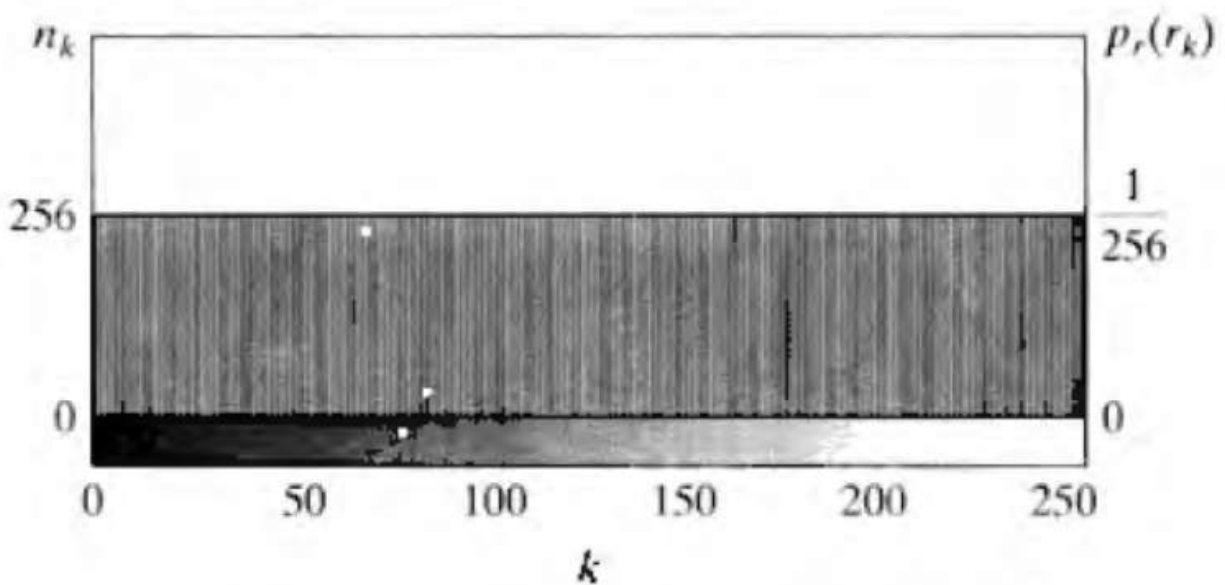
## 2. 空间冗余和时间冗余

考虑下图所示的计算机生成的恒定灰度线集合：



在对应的二维灰度阵列 中：

- 256中灰度等概率，如下直方图
- 每条线随机选择，所以垂直方向上，每条线的像素彼此独立
- 水平方向上则完全互相依赖（最大相关）



此时变长编码已经不能用于压缩该图像，可以基于行程长度指定一个新灰度的开始点和具有该灰度的连续像素的数量，基于行程长度的表示用  $(256 \times 256 \times 8) / [(256 + 256) \times 8]$  （128 : 1）的比率压缩图像\$。

多数情况下，像素在空间和时间上是相关的，可根据相邻像素对其进行合理预测，则它的视觉贡献多数情况下是冗余的。

### 3. 不相关信息

考虑一张每个像素灰度值都接近（视觉系统上看不出差异）的图像，那么它可以用达到  $256 \times 256 : 1$  的压缩比，但是去除多余的信息会导致定量信息的损失，这种信息的去除称为量化，是一种不可逆的操作。

### 4. 图像信息的度量

将信息的产生用一个概率过程建模，一个具有概率  $P(E)$  的随机事件  $E$  可被说成是包含  $I(E) = \log \frac{1}{P(E)} = -\log P(E)$  单位的信息。其中如果以  $m$  为对数的底，则度量称为  $m$  元单位，如果底为 2，则信息单位为比特，每个信源输出的平均信息称为该信息的熵  $H = -\sum_{j=1}^J P(a_j) \log P(a_j)$  ( $a_j$  信源符号，若彼此独立，则零记忆信源)，将一幅图考虑为一个虚构的零记忆灰度信源的输出，则灰度信源的熵变为  $H = \sum_{k=0}^{L-1} p_r(r_k) \log_2 p_r(r_k)$ ，上文提到的图像（边长编码计算那张）的熵可以代入式子进行计算：

$$\begin{aligned}\tilde{H} &= -[0.25 \log_2 0.25 + 0.47 \log_2 0.47 + 0.25 \log_2 0.25 + 0.03 \log_2 0.03] \\ &\approx -[0.25(-2) + 0.47(-1.09) + 0.25(-2) + 0.03(-5.06)] \\ &\approx 1.6614 \text{ 比特/像素}\end{aligned}$$

离题：当信息源的输出依赖于前面有限数量的输出时，那么信源称为马尔可夫信源或有限记忆信源。

### 5. 保真度准则

令  $\hat{f}(x, y)$  是  $f(x, y)$  的近似（图像压缩又解压），二者的误差  $e(x, y) = \hat{f}(x, y) - f(x, y)$ ，则总误差  $\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]$ ，则均方根误差为：

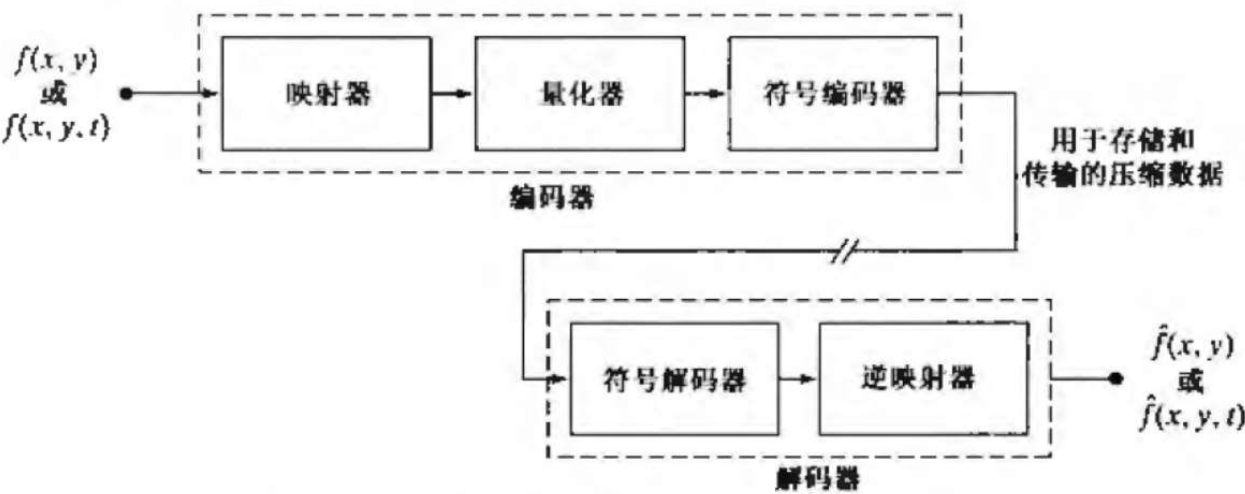
$e_{rms} = \left[ \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]^2 \right]^{1/2}$ ，若将  $\hat{f}(x, y)$  视作原始图像  $f(x, y)$  和一个误差或“噪声”信号的和，那么可用  $SNR_{rms}$  表示输出图像的均方信噪比如下：

$$SNR_{ms} = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \hat{f}(x, y)^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]^2}$$

但这只是一个客观的标准，图像的质量归根结底还是人来看的，所以有时候会出现客观标准差但人眼视觉看好的情况。

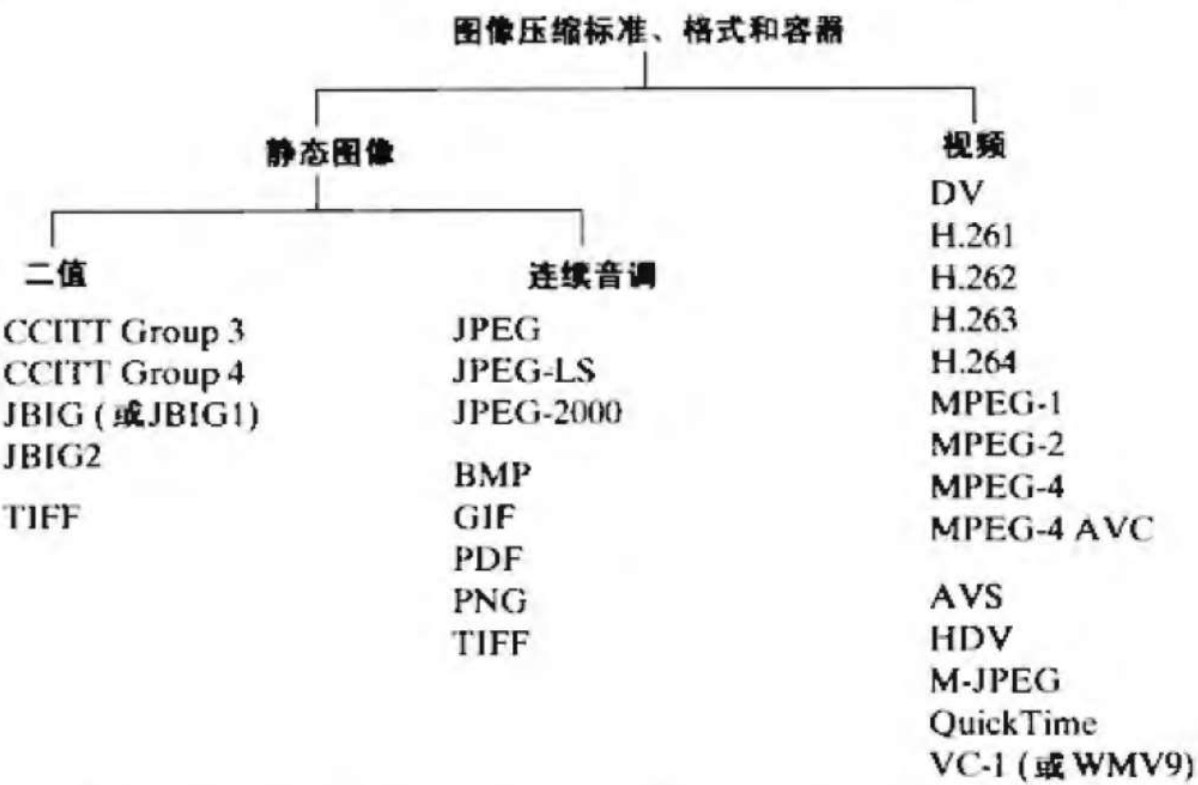
### 6. 图像压缩模型

一个编解码模型，编码模块执行压缩操作，解码模块执行解压缩的互补操作。如果重建的图像是精确复制品，则压缩系统称为无误差的、无损的或信息保持的压缩系统，如果不是，重建输出图像就会失真，称为有损压缩系统，如下所示，其中因为量化是不可逆的，所以解码器没有对应的操作。



其中映射器降低空间/时间冗余，量化器排除无关信息，编码器进行编码，通常是变长编码。解码器与之对应，操作相反。

7. 压缩标准



**二值静止图像**

CCITT 3 组	ITU-T	为在电话线上传输二值文件的传真(FAX)设计的方法。支持一维和二维行程编码[8.2.5]及霍夫曼编码[8.2.1]
CCITT 4 组	ITU-T	CCITT 3 组标准的一个精简和革新的版本, 仅支持二维行程编码
JBIG 或 JBIG1	ISO/IEC/ ITU-T	用于渐进、无损压缩二值图像的一个联合二值图像专家组标准。高达 6 比特/像素的连续色调图像可在一个位面的基础上编码[8.2.7]。使用了上下文敏感的算术编码[8.2.3], 且初始低分辨率的图像版本可使用附加的压缩数据逐步增强
JBIG2	ISO/IEC/ ITU-T	遵循 JBIG1 的, 针对桌面、互联网和 FAX 应用的二值图像的标准。使用的压缩方法是基于内容的: 对文本和半色调区域使用基于字典的方法[8.2.6], 而对其他图像内容使用霍夫曼编码[8.2.1]或算术编码[8.2.3]。它可以是有损的或无损的

**连续色调静止图像**

JPEG	ISO/IEC/ ITU-T	针对图片质量的图像的联合图片专家组标准。其有损基准编码系统(最通用的实现)在 $8 \times 8$ 图像块上[8.2.8]用量化的离散余弦变换(DCT)、霍夫曼编码[8.2.1]和行程编码[8.2.5]。它是在互联网上压缩图像的最普遍的方法之一
------	-------------------	--

**连续色调静止图像**

JPEG-LS	ISO/IEC/ ITU-T	针对连续色调图像的基于自适应预测[8.2.9]、上下文模型[8.2.3]和 Golomb 编码[8.2.2]的无损到接近无损的标准
JPEG-2000	ISO/IEC/ ITU-T	遵循 JPEG 的用于提高图片质量图像的压缩率的标准。它使用算术编码[8.2.3]和量化离散小波变换(DWT)[8.2.10]。压缩可以是有损的或无损的

**视频**

DV	IEC	数字视频。适用于家庭和半专业视频制作应用, 以及类似电子新闻采集和便携式摄像机等设备的视频标准。它使用一种类似于 JPEG 的基于 DCT 的方法[8.2.8]对简单编辑目的单独地压缩图像帧
H.261	ITU-T	针对 ISDN 线路(综合业务数字网)的一种双向视频会议标准。它分别支持非隔行扫描的 $352 \times 288$ 和 $176 \times 144$ 分辨率图像[称为 CIF(通用中间格式)和 QCIF(四分之一 CIF)]。采用类似于 JPEG 的基于 DCT 的压缩方法[8.2.8], 使用帧间预测差分[8.2.9]来减少时间冗余。基于块的技术用于帧间的运动补偿
H.262	ITU-T	见下面的 MPEG-2
H.263	ITU-T	针对具有附加分辨率 SQCIF(Sub-Quarter CIF $128 \times 96$ )、4CIF( $704 \times 576$ )和 16CIF( $1408 \times 512$ )的普通电话调制解调器(如 28.8 Kbps)设计的 H.261 的增强版本
H.264	ITU-T	针对视频会议、网络流和电视广播的 H.261 至 H.263 的一个扩展。它支持帧间预测差分[8.2.9]、可变块大小的整数变换(不是 DCT)和上下文自适应算术编码[8.2.3]
MPEG-1	ISO/IEC	针对 CD-ROM 应用的非隔行扫描视频的速率高达 1.5 Mb/s 的运动图像专家组标准。它类似于 H.261, 但帧预测可基于前一帧、下一帧或两者的内插。几乎所有的计算机和 DVD 播放器都支持该标准
MPEG-2	ISO/IEC	针对 DVD 所设计的传输速率高达 15 Mb/s 的 MPEG-1 的一个扩展。支持隔行扫描视频和 HDTV。它是目前为止最为成功的视频标准
MPEG-4	ISO/IET	MPEG-2 的一个扩展, 它支持可变块大小和帧间预测差分[8.2.9]
MPEG-4AVC	ISO/IET	MPEG-4 的第 10 部分, 即先进视频编码(AVC)。与上面的 H.264 AVC 相同

**连续色调静止图像**

BMP	Microsoft	Windows 位图。主要用做未压缩的简单图像的一种文件格式
GIF	CompuServe	图形交换格式。对 1 到 8 比特图像使用无损 LZW 编码[8.2.4]的文件格式。它常用于为万维网制作小动画和低分辨率影片
PDF	Adobe Systems	便携文件格式。针对与设备和分辨率无关的方法表示二维文档的一种格式。它可以像 JPEG、JPEG-2000、CCITT 和其他压缩图像的容器来运行。某些 PDF 版本已成为 ISO 标准
PNG	万维网联盟(W3C)	便携网络图形。通过对每个像素值与基于过去像素的预测值间的差进行编码, 是全透明(高达 48 比特/像素)无损地压缩全彩色图像[8.2.9]的文件标准。
TIFF	Aldus	标记图像格式。一种灵活的文件格式, 支持多种图像压缩标准, 包括 JPEG、JPEG-LS、JPEG-2000、JBIG2 和其他标准

**视频**

AVS	MII	音视频标准。类似于 H.264, 但使用指数 Golomb 编码[8.2.2]。中国开发的标准
HDTV	公司联盟	高清晰度视频。针对高清电视的 DV 的一个扩展, 它使用类似于 MPEG-2 的压缩, 包括使用预测差分方法[8.2.9]去除时间冗余
M-JPEG	多家公司	活动 JPEG。一种压缩格式, 在这种格式中, 每一帧都独立地用 JPEG 来压缩
Quick-Time	Apple Computer	支持 DV、H.261、H.262、H.264、MPEG-1、MPEG-2 和 MPEG-4 他视频压缩格式的一种媒体容器
VC-1	SMPTE	互联网上最为通用的视频格式
WMV9	Microsoft	适用于高清和蓝光高分辨率 DVD 的一种格式。它类似于 H.264/AVS, 使用一个可变块大小的整数 DCT[8.2.8, 8.2.9]和依赖于上下文的变长码表[8.2.1], 但没有帧间预测

---

欢迎扫描二维码关注微信公众号 深度学习与数学 [每天获取免费的大数据、AI等相关的学习资源、经典和最新的深度学习相关的论文研读, 算法和其他互联网技能的学习, 概率论、线性代数等高等数学知识的回顾]

