第八章 图像编码

主要内容：

1.图像编码的必要性

2.图像编码的分类

3.图像编码中的保真度准则

4.编码的性能参数

5.编码冗余

6.相关性冗余

7.图像编码的国际标准

上海大学
Shanghai University

8.1 图像压缩和数据冗余

（1）图像编码的必要性

- 图像编码与压缩

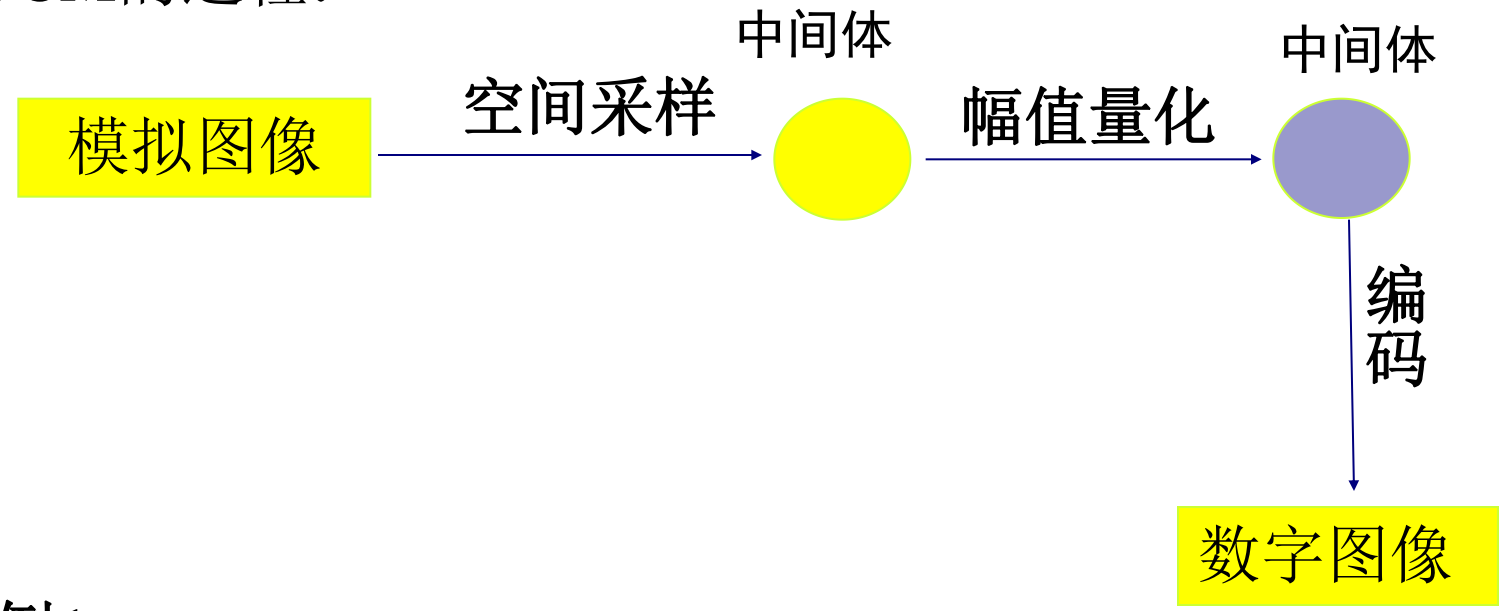
图像编码与压缩，本质上来说，就是对图像源数据按一定的规则进行变换和组合，从而达到以尽可能少的代码来表示尽可能多的数据信息.压缩通过编码来实现，或者说编码带来压缩的效果，所以，一般把此项处理称之为压缩编码.

- 编码的必要性

一幅模拟图像必须经过脉码调制（PCM—Pulse Code Modulation)才能变成数字图像.（PCM有时也指对信号进行采样、量化并以适当码字将其编码的各个过程的总称）

http://mdedu.bbi.edu.cn/resource/courses/html/dmtjs/content/study/3-0.htm#

PCM的过程：



例1

设一幅活动图像的空间分辨率为N,灰度分辨率为b, 时间分辨率为f_B, 则在实时传输过程中，该图像在传输通道里的传输率至少应该为

$$\rho=Nbf_B$$

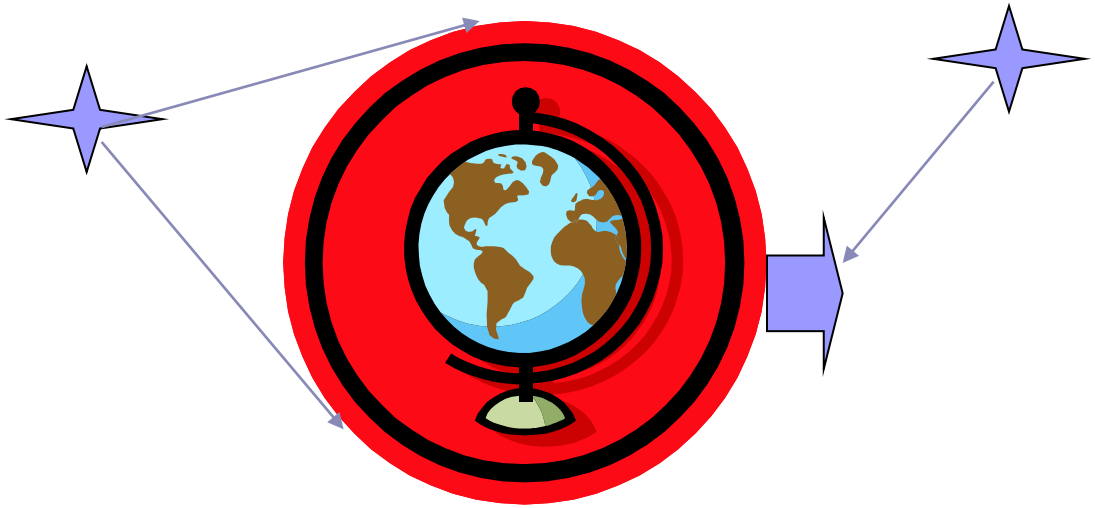
若N=512×512, b=8, f_B=25, 则ρ=52.4Mbps.

例2

地球资源卫星(LANDSAT)一帧图像(4幅)的数据量为

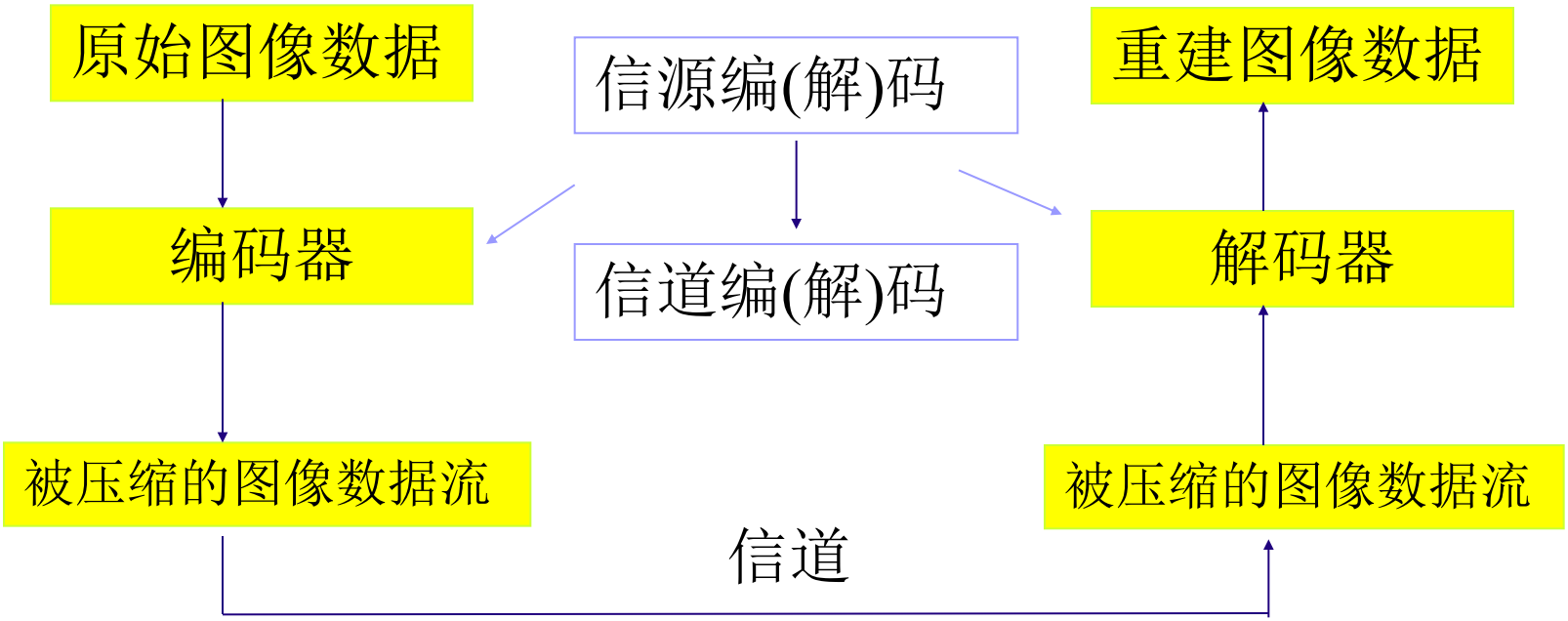
$$2340 \times 2340 \times 7 \times 4 = 153{,}316{,}800 \cong 153\text{Mb}$$

卫星每天要获取很多幅图像，这些数据都先暂时存储在卫星体内的磁性存储器中，当卫星飞过地面接收站的有效接收区域时，迅速将这些数据全部送到地面.



图像编码的目的：节省存储空间；减少传输时间；利于处理，降低处理成本.

图像数据经过编码压缩、传输、解码以及重建图像数据的流程如下图所示：



信息保持型编码　　信息损失型编码

➢压缩率

$$C_R = \frac{n_1}{n_2}$$

n1:第一种编码方案下的数据量;

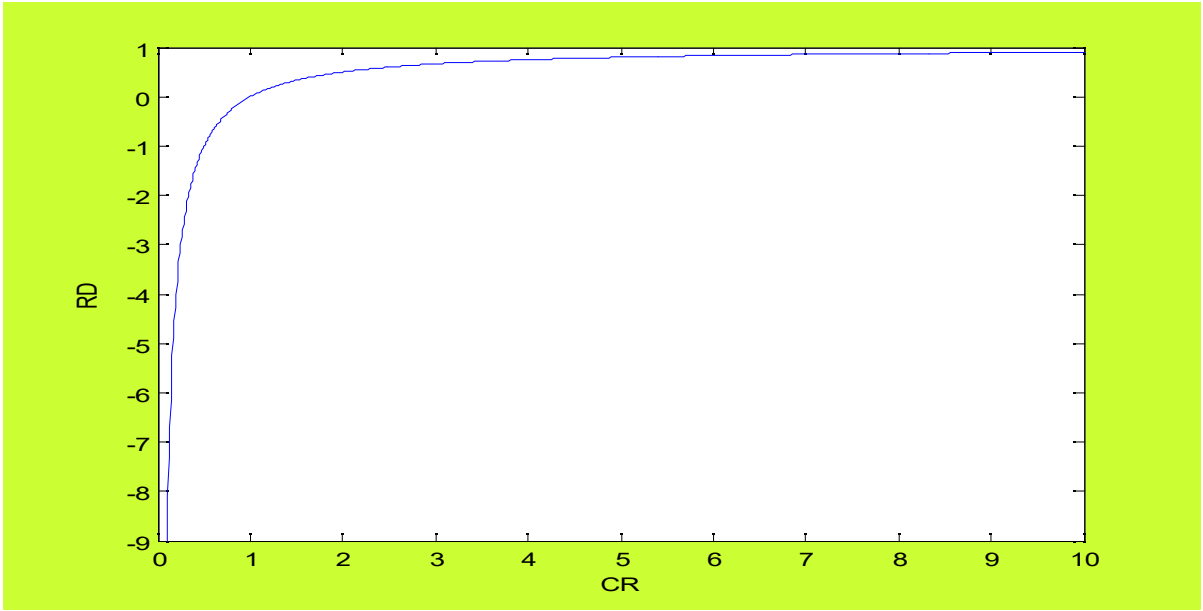
n2:第二种编码方案下的数据量.

➢相对数据冗余度

$$R_D = 1 - \frac{1}{C_R}$$

$$= \frac{n_1 - n_2}{n_1}$$

$R_D = 1 - \frac{1}{C_R}$ 的函数图像



第一种编码方案下的数据集合相对于第二种编码方案下的数据集合的冗余量.

（2）图像编码压缩分类

a)从应用角度分类

静止图像编码，活动图像编码，二值图像编码

b)从信息保持程度角度分类

有损压缩（保真度编码，特征抽取编码）

无损压缩（信息保持压缩，熵保持压缩）

c)从具体的编码技术角度分类

空域法，变换域法

预测编码，变换编码，统计编码，等

（3）图像数据的冗余

冗余大致分为三大类

1）编码冗余（也称为信息熵冗余）

符号序列⇒码字⇒（码字长度）

2）像素间相关性冗余

帧间像素信息冗余（时间冗余），帧内像素信息冗余（空间冗余）

3）视觉冗余

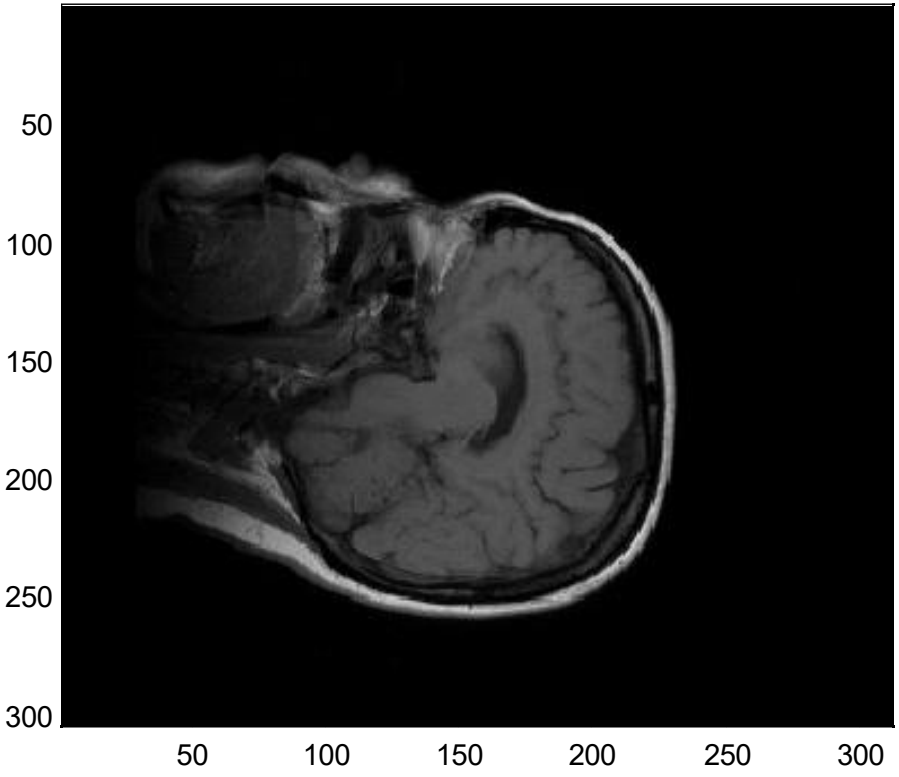
人眼对所有视觉信息并不是都具有相同的敏感度；

人眼的空间分辨率，时间分辨率.

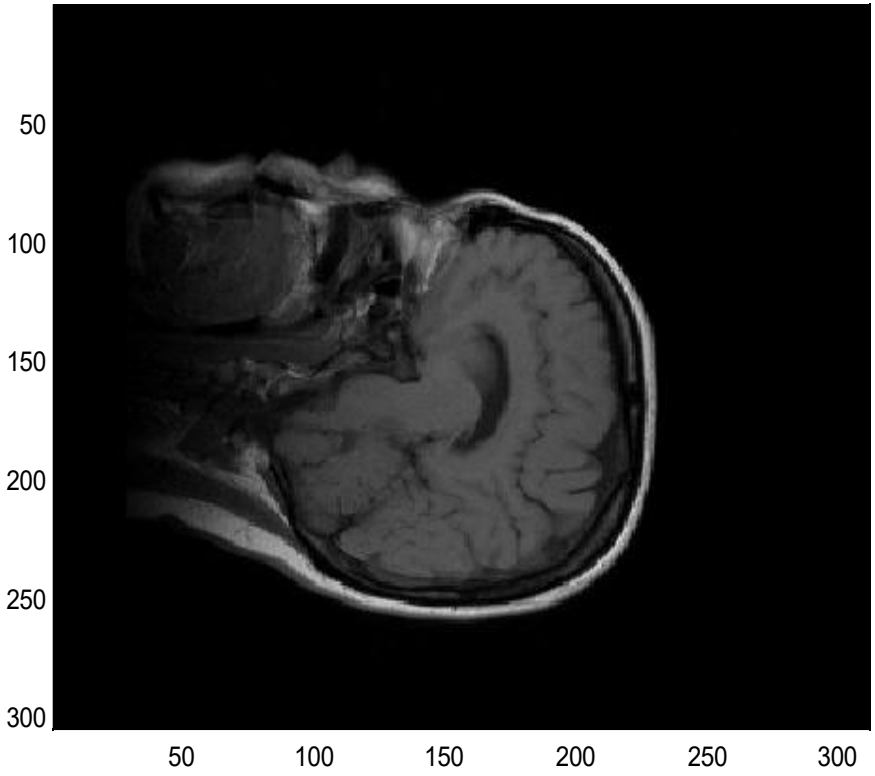
消除冗余能达到数据压缩的效果.

例3 编码冗余

```
a=imread('2.bmp');
b= uint16(a);
 imagesc(b*255);
```

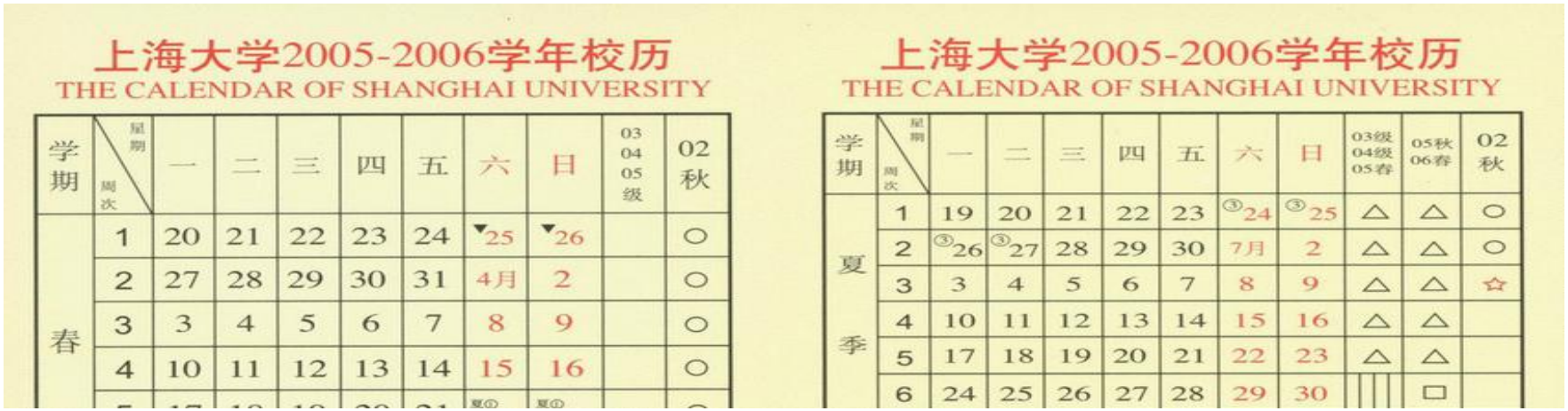



```
a=imread('2.bmp');
c= uint8(a);
 imagesc(c);
```



例4 相关性冗余

帧内相关性



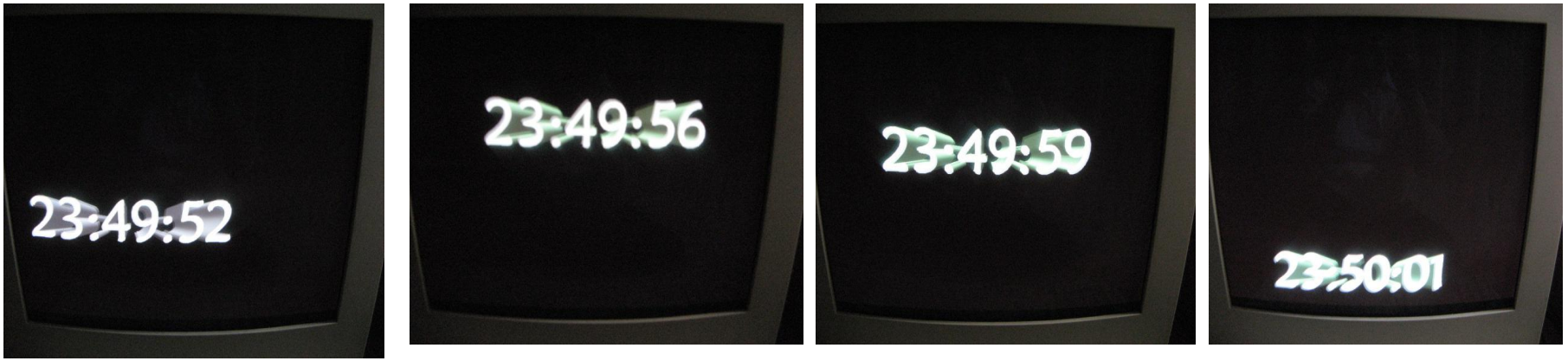
上海大学2005-2006学年校历
THE CALENDAR OF SHANGHAI UNIVERSITY

学期	星期 / 周次	一	二	三	四	五	六	日	03 04 05 级	02 秋
春	1	20	21	22	23	24	▼25	▼26		○
	2	27	28	29	30	31	4月	2		○
	3	3	4	5	6	7	8	9		○
	4	10	11	12	13	14	15	16		○

上海大学2005-2006学年校历
THE CALENDAR OF SHANGHAI UNIVERSITY

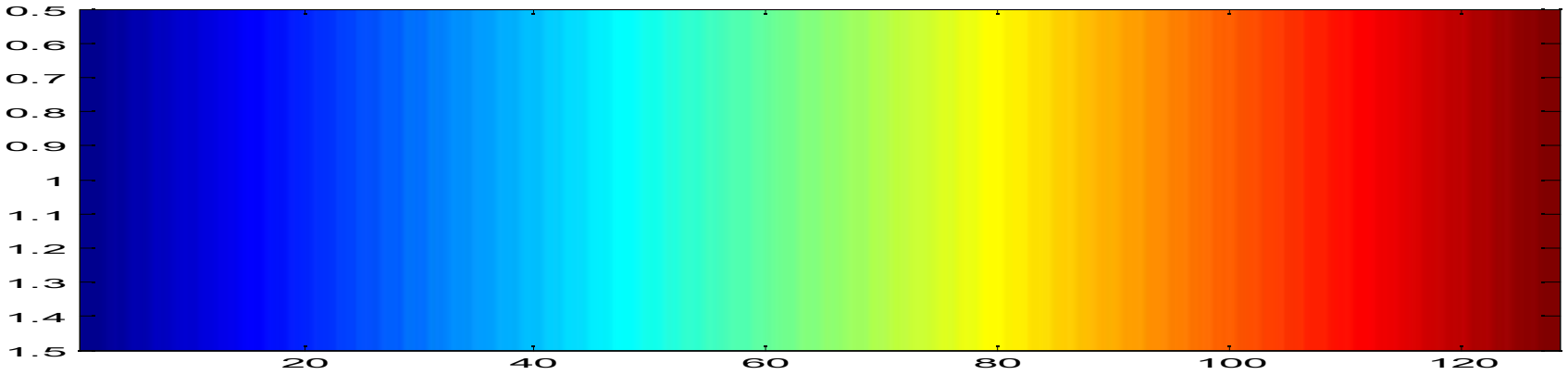
学期	星期 / 周次	一	二	三	四	五	六	日	03级 04级 05春	05秋 06春	02 秋
夏季	1	19	20	21	22	23	③24	③25	△	△	○
	2	③26	③27	28	29	30	7月	2	△	△	○
	3	3	4	5	6	7	8	9	△	△	☆
	4	10	11	12	13	14	15	16	△	△	
	5	17	18	19	20	21	22	23	△	△	
	6	24	25	26	27	28	29	30	‖	□	

帧间相关性

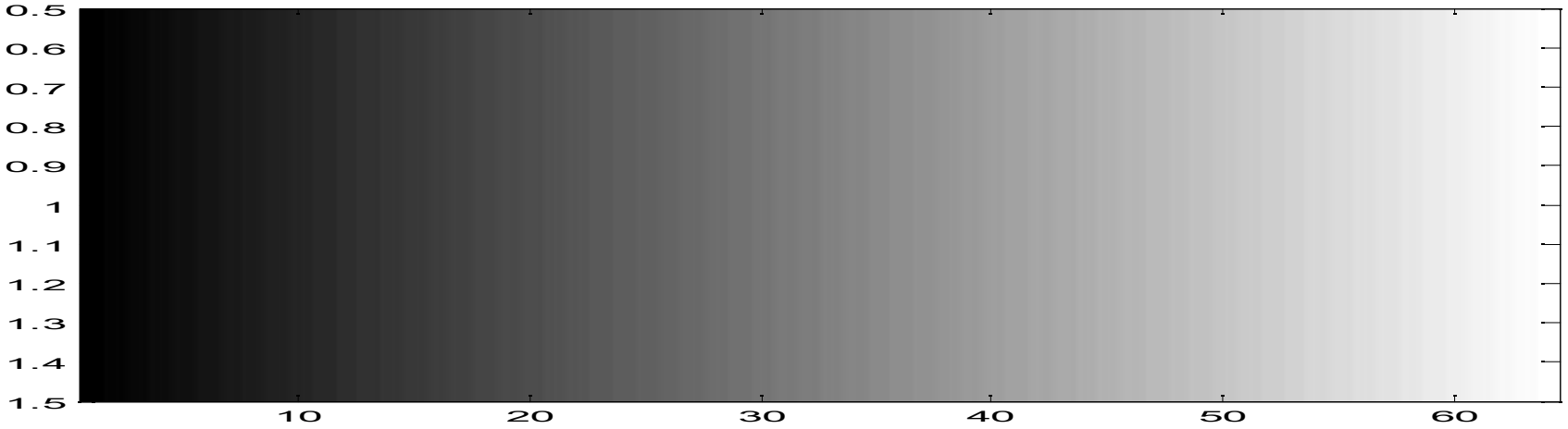


例5 视觉冗余

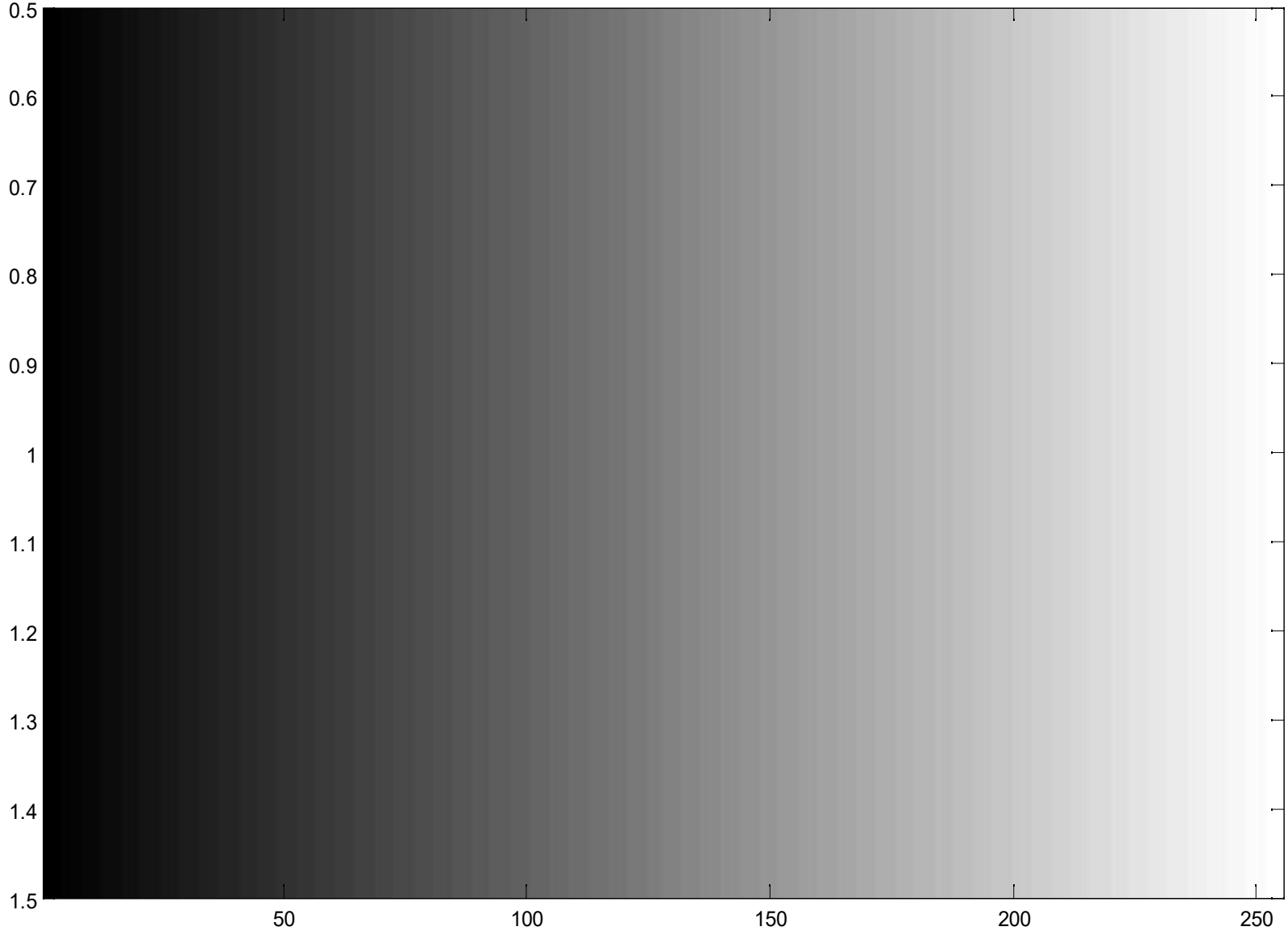
colormap(jet); imagesc(1:128);



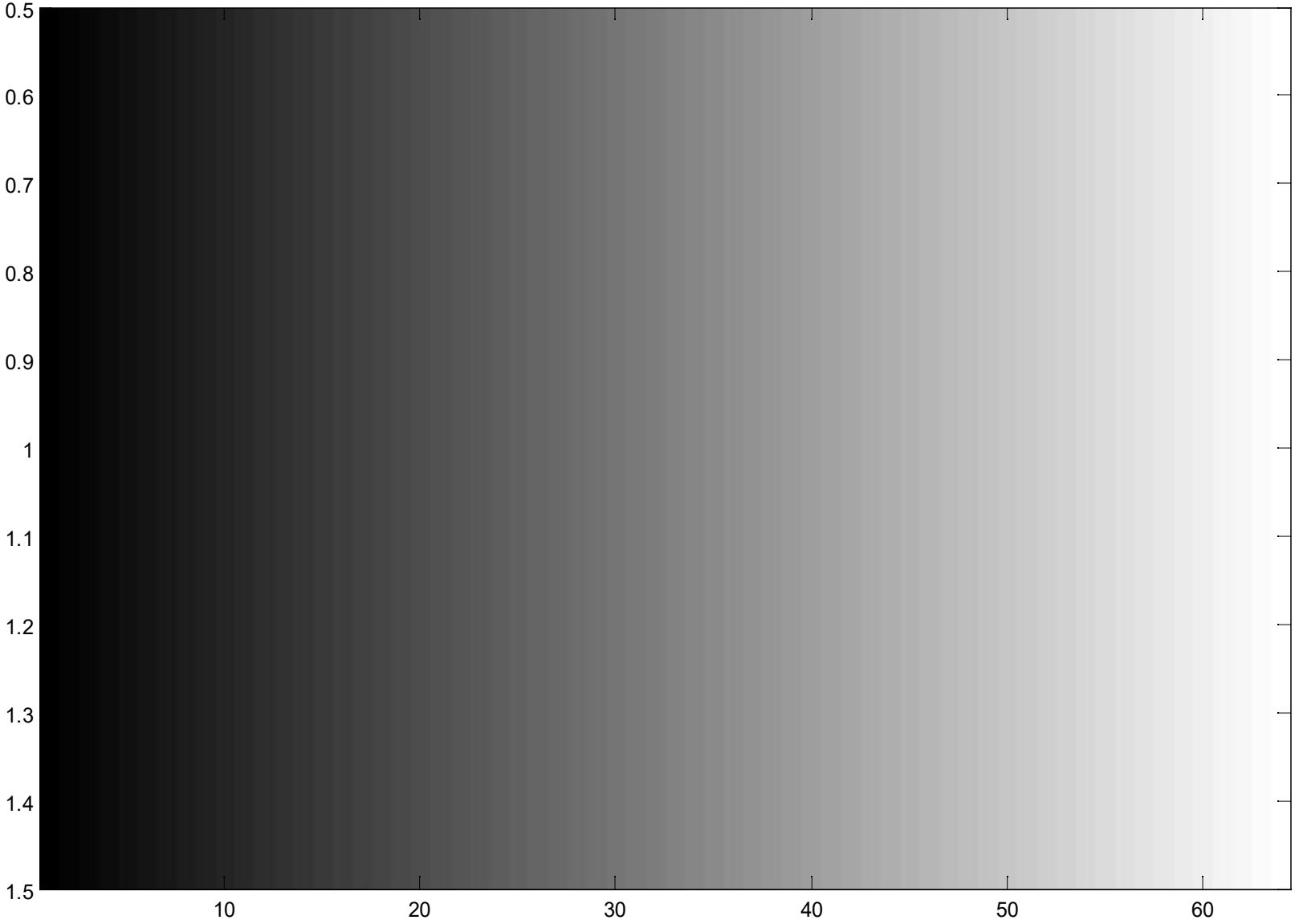
colormap(gray); imagesc(1:64);



colormap(gray); imagesc(1:256);



colormap(gray); imagesc(1:64);



（4）图像的编码质量评价

图像品质的核心问题是逼真度问题.经过处理的图像（包括经过压缩编码后的图像）与一个标准图像之间的偏差可以作为图像逼真度（保真度）的度量.这一偏差，包括亮度，色度，分辨率以及某些心理物理学参数.

1)客观评价准则

设f(x,y)是输入图像，f’(x,y)是输出图像, 定义偏差e(x,y)=f(x,y)-f’(x,y), 则以下的参数可作为保真度准则：

总偏差：$\sum_x \sum_y e(x,y)$

均方差：$\frac{1}{N^2} \sum_x \sum_y e^2(x,y)$

均方信噪比：$\sum_x \sum_y f'^2(x,y) / \sum_x \sum_y e^2(x,y)$

设$\bar{f}=\sum_x\sum_y f(x,y)$，基本信噪比：

$$10\lg\left(\sum_x\sum_y (f(x,y)-\bar{f})^2 / \sum_x\sum_y e^2(x,y)\right)$$

2)主观评价准则

1	2	3	4	5	6	7
很差	较差	稍差	相同	稍好	较好	很好

（5） 编码定理

问题：如何度量编码方法的优劣？（编码的性能参数）

➢图像信息熵与平均码字长度

令 $d=\{d_1,d_2,\cdots,d_m\}$ 是图像像素灰度级集合

其对应的频率为 $p(d_1),p(d_2),\cdots,p(d_m)$ 定义

$$H(d)=-\sum_{i=1}^{m}p(d_i)\log_2 p(d_i)$$ （单位：比特/像素）

令 $\{\beta_1,\beta_2,\cdots,\beta_m\}$ 是对应像素灰度级的编码长度，定义

$$R(d)=\sum_{i=1}^{m}p(d_i)\beta_i$$ （单位：比特/像素）

称H(d)为该图像的平均信息熵，R(d)为平均编码长度.

- 编码效率

$$\eta = \frac{H(d)}{R(d)}(\%)$$

- 冗余度

$$R_d = 1 - \eta$$

➢熵与平均码字长度

1）H(d)<<R(d)时，一定可以设计出某种平均码字长更短的无失真编码方法.

2）平均码字长小于H(d)的无失真编码方法不存在.

➢熵编码

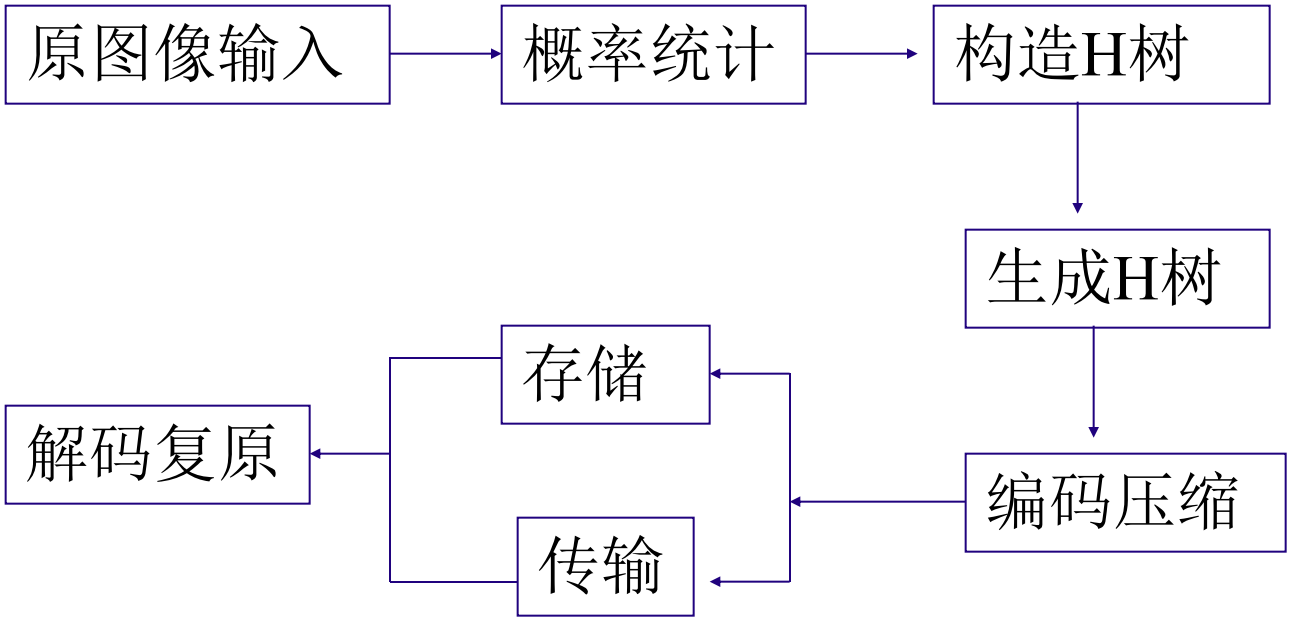
使编码后的图像的平均码字长度尽可能接近图像的熵H.

基本思路是：概率大的灰度级用短码字，概率小的，用长码字.

根据图像像素灰度值出现的概率的分布特性而进行的压缩编码叫统计编码.

8.2 一些基本的压缩方法

➢Huffman编码



基本哈夫曼编码系统框图

去除编码冗余

❖算法

1）将灰度等级按概率大小进行排序（降序），

每个灰度等级作为一个叶子结点，形成一棵树；

2）将两个根节点概率最小的树，合并（规则：这

两个结点构造一个双亲结点，双亲结点的概率

大小是两者之和）；重复1) 2), 直到只有一个树为止；

3）设所有左后代为0，右后代为1.

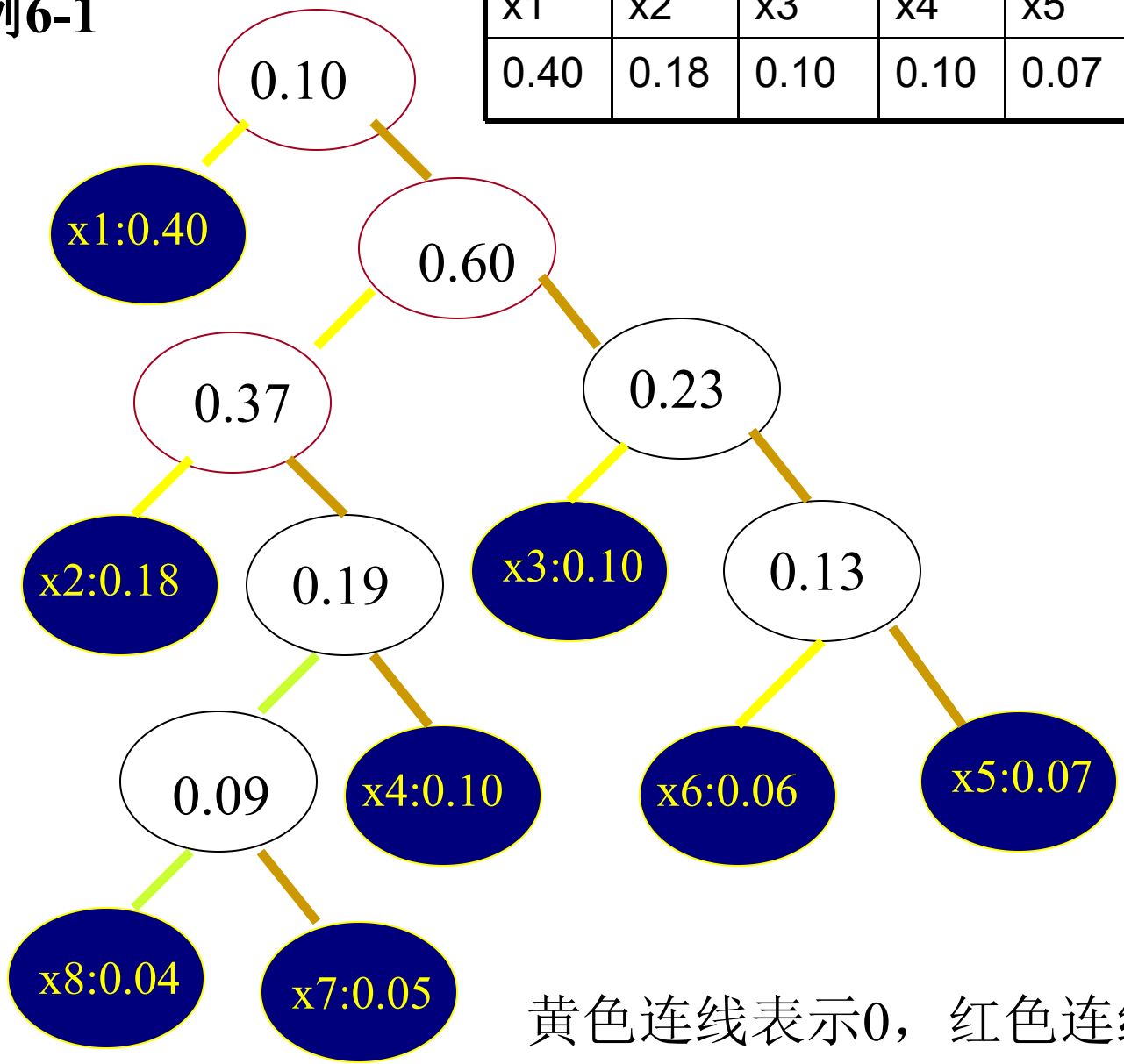
❖特点

优点：即时码；最优码

缺点：当需要对大量符号进行编码时，构造最优哈夫曼码的计算量会很大.

例6-1

x1	x2	x3	x4	x5	x6	x7	x8
0.40	0.18	0.10	0.10	0.07	0.06	0.05	0.04



x1:0

x2:100

x3:110

x4:1011

x5:1111

x6:1110

x7:10101

x8:10100

黄色连线表示0，红色连线表示1

信源符号N,则信源消减次数为N-2(子树合并次数）

平均码长：

$$R(d)=\sum_{i=1}^{8}p(d_i)\beta_i$$
$$=0.40\times1+0.18\times3+0.10\times3+0.10\times4$$
$$+0.07\times4+0.06\times4+0.05\times5+0.04\times5$$
$$=2.61$$

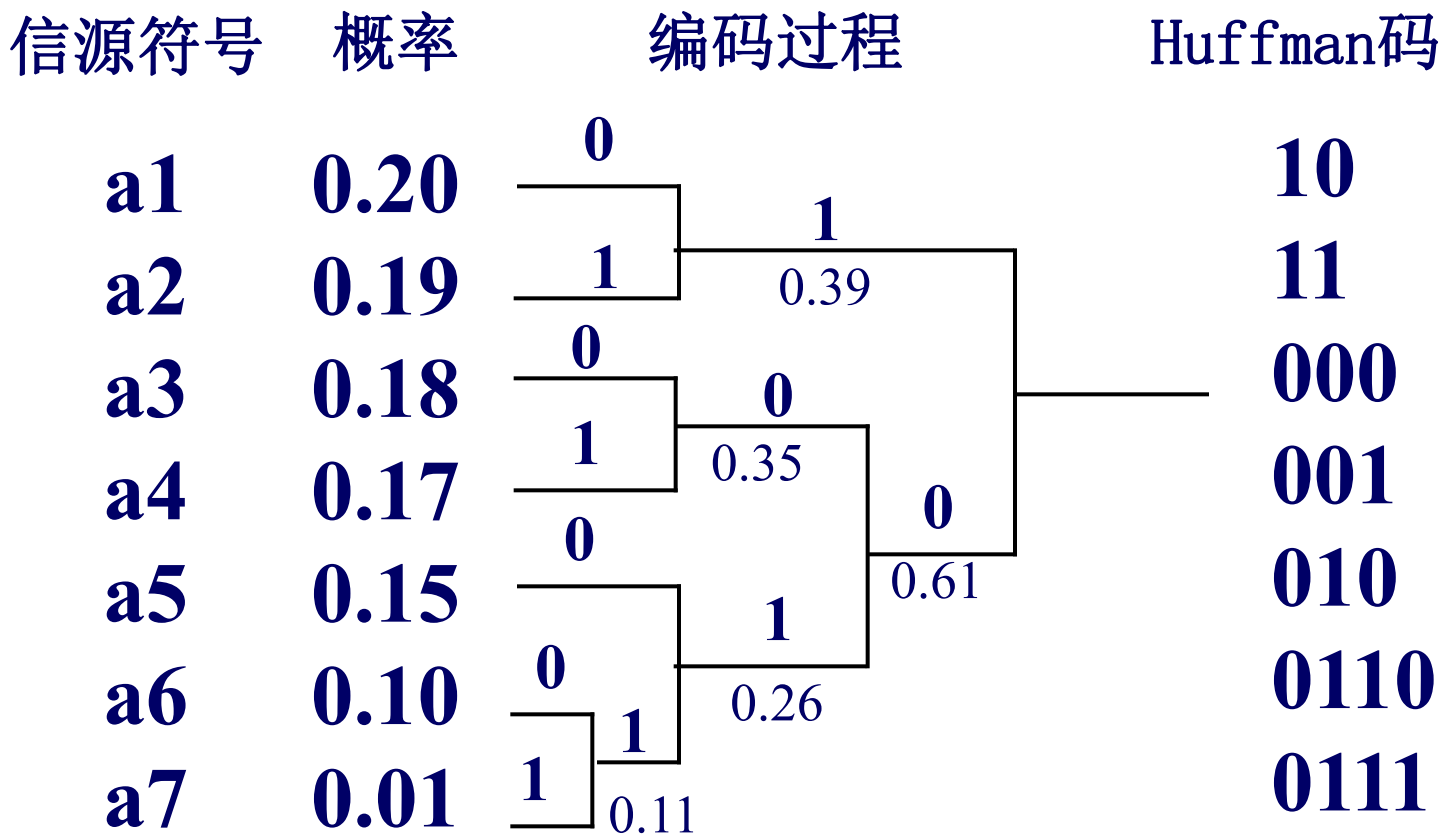
信息熵：

$$H(d)=-\sum_{i=1}^{8}p(d_i)\log_2 p(d_i)$$
$$=0.40\log_2 0.40+0.18\log_2 0.18+0.10\log_2 0.10+0.10\log_2 0.10$$
$$+0.07\log_2 0.07+0.06\log_2 0.06+0.05\log_2 0.05+0.04\log_2 0.04$$
$$=2.55$$

编码效率：

$$\eta=\frac{H(d)}{R(d)}(\%)=2.25/2.61=97.8\%$$

例6-2 http://www.tvdiy.net/tv/6609-16.htm



Huffman编码过程

哈夫曼码的改型

表 9.4.2　哈夫曼码与其改型的比较

块号	信源符号	概率	截断哈夫曼码		平移哈夫曼码		哈夫曼码
第 1 块	b_1	0.25	01	01	01	01	01
	b_2	0.21	10	10	10	10	10
	b_3	0.19	000	11	000	11	11
	b_4	0.16	001	001	001	001	001
第 2 块	b_5	0.08	11 00	000 00	11 01	000 01	0001
	b_6	0.06	11 01	000 01	11 10	000 10	00000
	b_7	0.03	11 10	000 10	11 000	000 11	000010
	b_8	0.02	11 11	000 11	11 011	000 001	000011
熵		2.65					
平均长度			2.73		2.78	2.75	2.7

亚最优编码方法，通过牺牲编码效率来换取编码计算量的减少.

截断哈夫曼码：对最可能出现的M个符号进行哈夫曼编码，对其他码都用一个合适的定长码之前加一个前缀码来表示.

➢Shannon-Fano编码

◆算法

1）将消息非递增排序

2）按概率之和相近或相等原则将消息集一分为二

3）将以上分割准则递归地应用到消息子集，直止 最终子集只有一个消息为止

4）在分割过程中，分别给所分得的两个子集赋予0和1

例7

<table>
<tr><td>x1</td><td>x2</td><td>x3</td><td>x4</td><td>x5</td><td>x6</td><td>x7</td><td>x8</td></tr>
<tr><td>0.40</td><td>0.18</td><td>0.10</td><td>0.10</td><td>0.07</td><td>0.06</td><td>0.05</td><td>0.04</td></tr>
<tr><td colspan="2">0</td><td colspan="6">1</td></tr>
<tr><td rowspan="4">0</td><td rowspan="4">1</td><td colspan="2">0</td><td colspan="4">1</td></tr>
<tr><td rowspan="3">0</td><td rowspan="3">1</td><td colspan="2">0</td><td colspan="2">1</td></tr>
<tr><td>0</td><td>1</td><td>0</td><td>1</td></tr>
</table>

x1:00

x2:01

x3:100

x4:101

x5:1100

x6:1101

x7:1110

x8:1111

平均码长：

$$R(d)=\sum_{i=1}^{8}p(d_i)\beta_i$$
$$=0.40\times2+0.18\times2+0.10\times3+0.10\times3$$
$$+0.07\times4+0.06\times4+0.05\times4+0.04\times4$$
$$=3.44$$

信息熵：

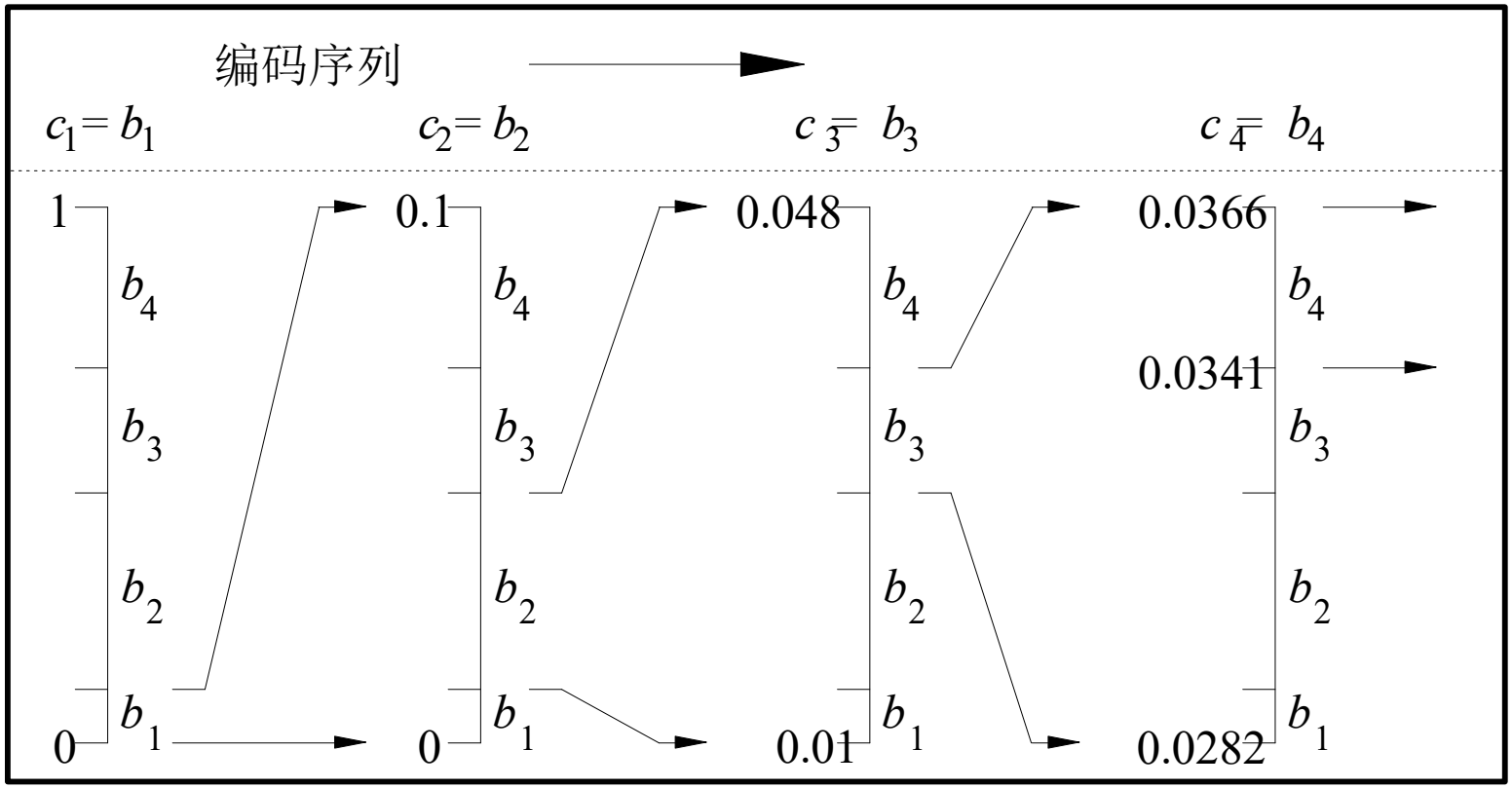
$$H(d)=-\sum_{i=1}^{8}p(d_i)\log_2 p(d_i)$$
$$=0.40\log_2 0.40+0.18\log_2 0.18+0.10\log_2 0.10+0.10\log_2 0.10$$
$$+0.07\log_2 0.07+0.06\log_2 0.06+0.05\log_2 0.05+0.04\log_2 0.04$$
$$=2.55$$

编码效率：

$$\eta=\frac{H(d)}{R(d)}(\%)=2.25/3.44=65.4\%$$

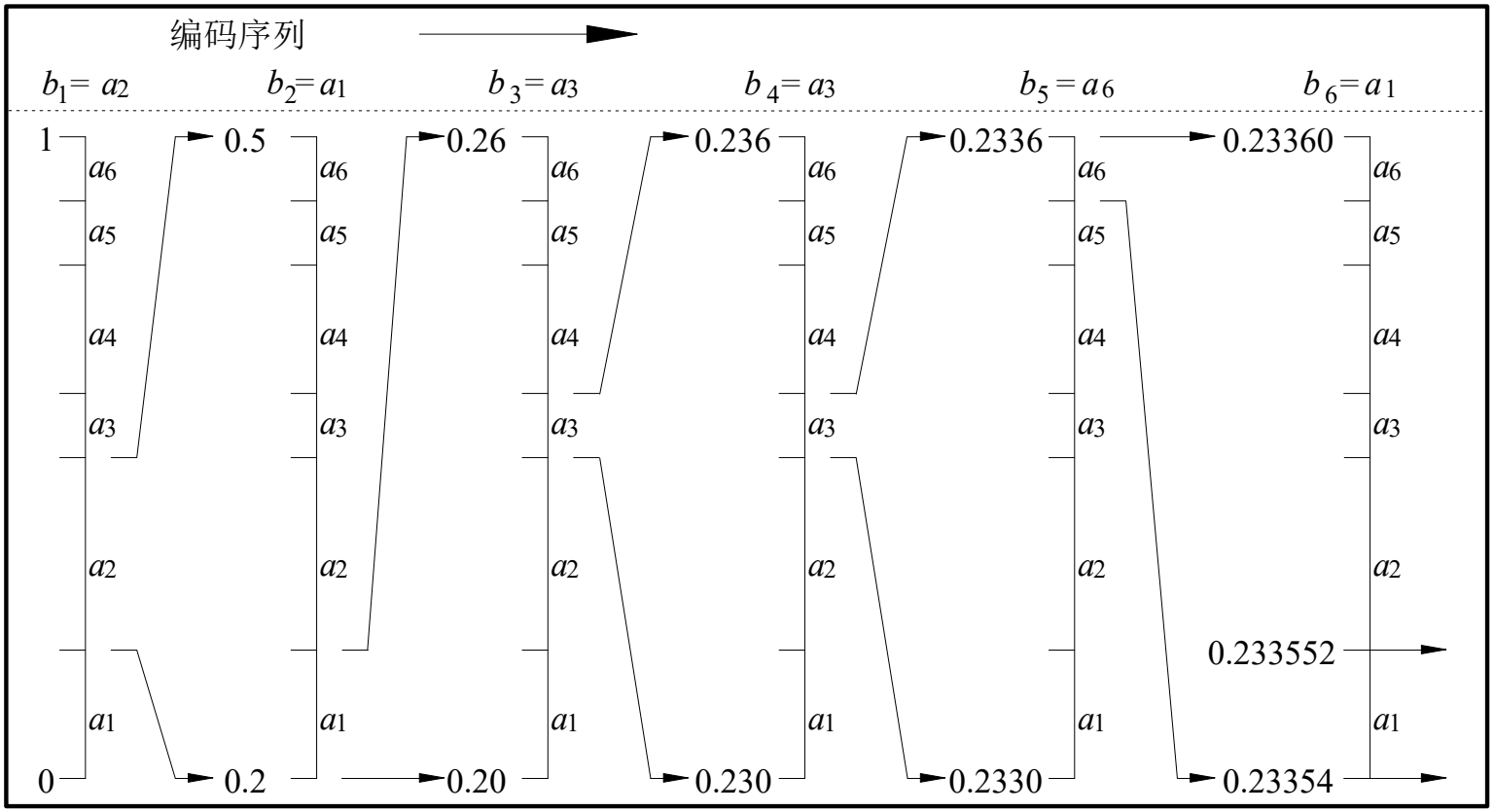
➢算术编码

算术编码过程图解(1)



信源符号和码字之间不存在一一对应关系.

去除编码冗余



算术编码的解码过程是借助对信源符号的编码过程进行的.

➢LZW编码（Lempel-Ziv-Welch)

也叫字典编码（美国的专利）

原理：将定长码字分配给变长信源符号序列。

算法：1）字典初始化，2）扫描像素，3）判断当前词是否是新单词，之后判断否有新单词出现？4）字典更新.

例：4×4图像

39	39	126	126
39	39	126	126
39	39	126	126
39	39	126	126

1)字典初始化

字典位置	条　目
0	0
1	1
⋮	⋮
255	255
256	—
⋮	⋮
511	—

表 8.7　LZW 编码示例

当前可识别的序列	正被处理的像素	编码后的输出	字典位置(码字)	字 典 词 条
	39			
39	39	39	256	39-39
39	126	39	257	39-126
126	126	126	258	126-126
126	39	126	259	126-39
39	39			
39-39	126	256	260	39-39-126
126	126			
126-126	39	258	261	126-126-39
39	39			
39-39	126			
39-39-126	126	260	262	39-39-126-126
126	39			
126-39	39	259	263	126-39-39
39	126			
39-126	126	257	264	39-126-126
126		126		

➢行程编码

1）扫描像素，2）记录（灰度，行程）对.

特别适合二值图像！

1	1	1	1	1	0	0	0	0	0
0	1	1	1	1	1	1	1	1	0
0	0	0	0	0	0	0	1	1	1
1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0	1	1

编码：(1)5,6,8,8,13,8,2

1	1	1	1	1	2	2	2	2	2
2	2	2	5	5	5	5	5	5	5
0	0	0	0	0	0	0	5	5	5
5	5	5	5	5	5	5	5	5	7
7	7	7	7	7	7	2	2	2	2

编码：(1,5)(2,8)(5,7)(0,7)(5,13)(7,7)(2,4)

去除空间冗余

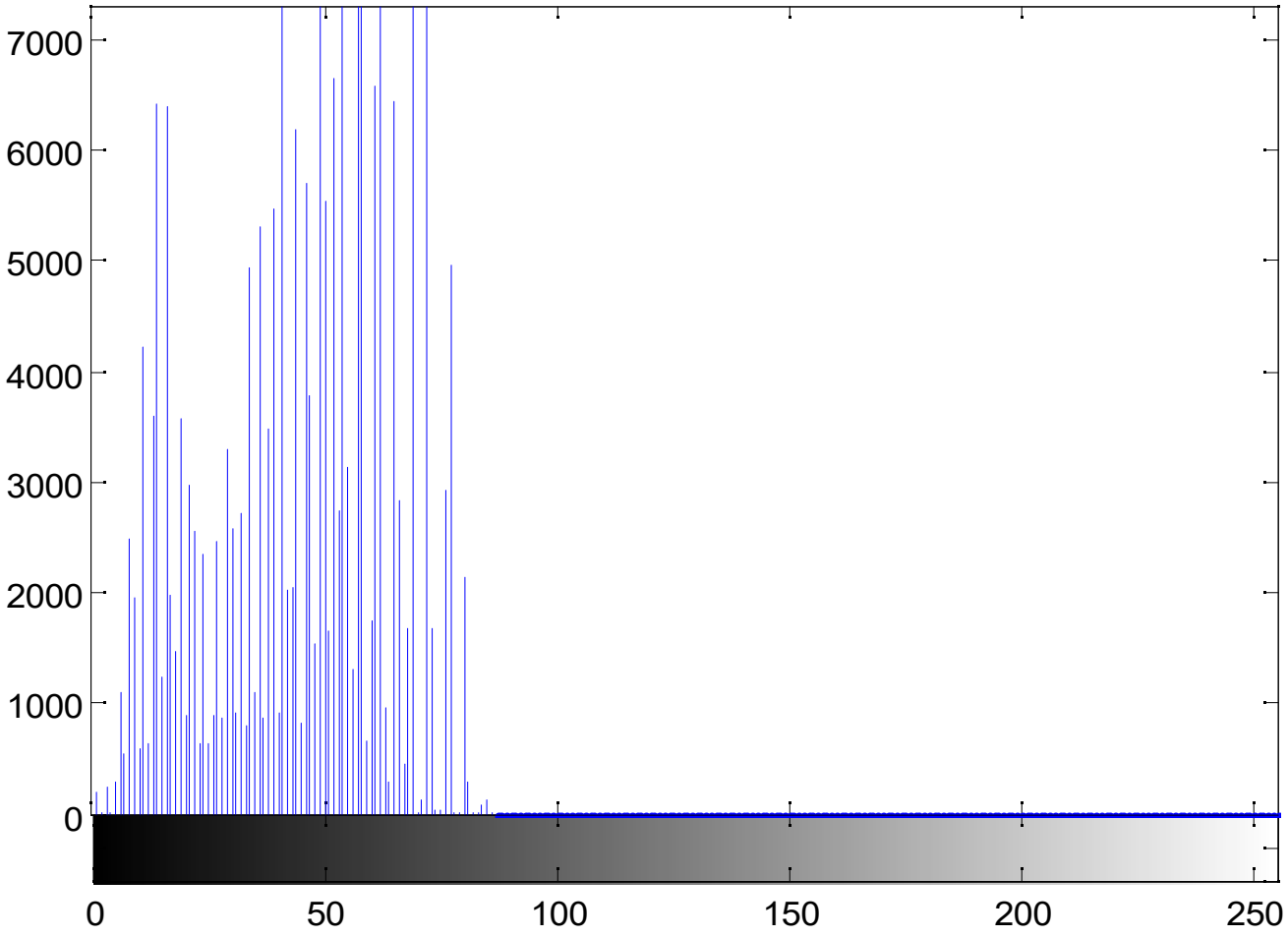
➢预测编码

预测编码（Predictive Coding)，就是根据“过去”的时刻的像素值，运用一种模型，预测当前的像素值，预测编码通常不直接对信号编码，而是对预测误差进行编码.当预测比较准确，误差较小时，即可达到编码压缩的目的.

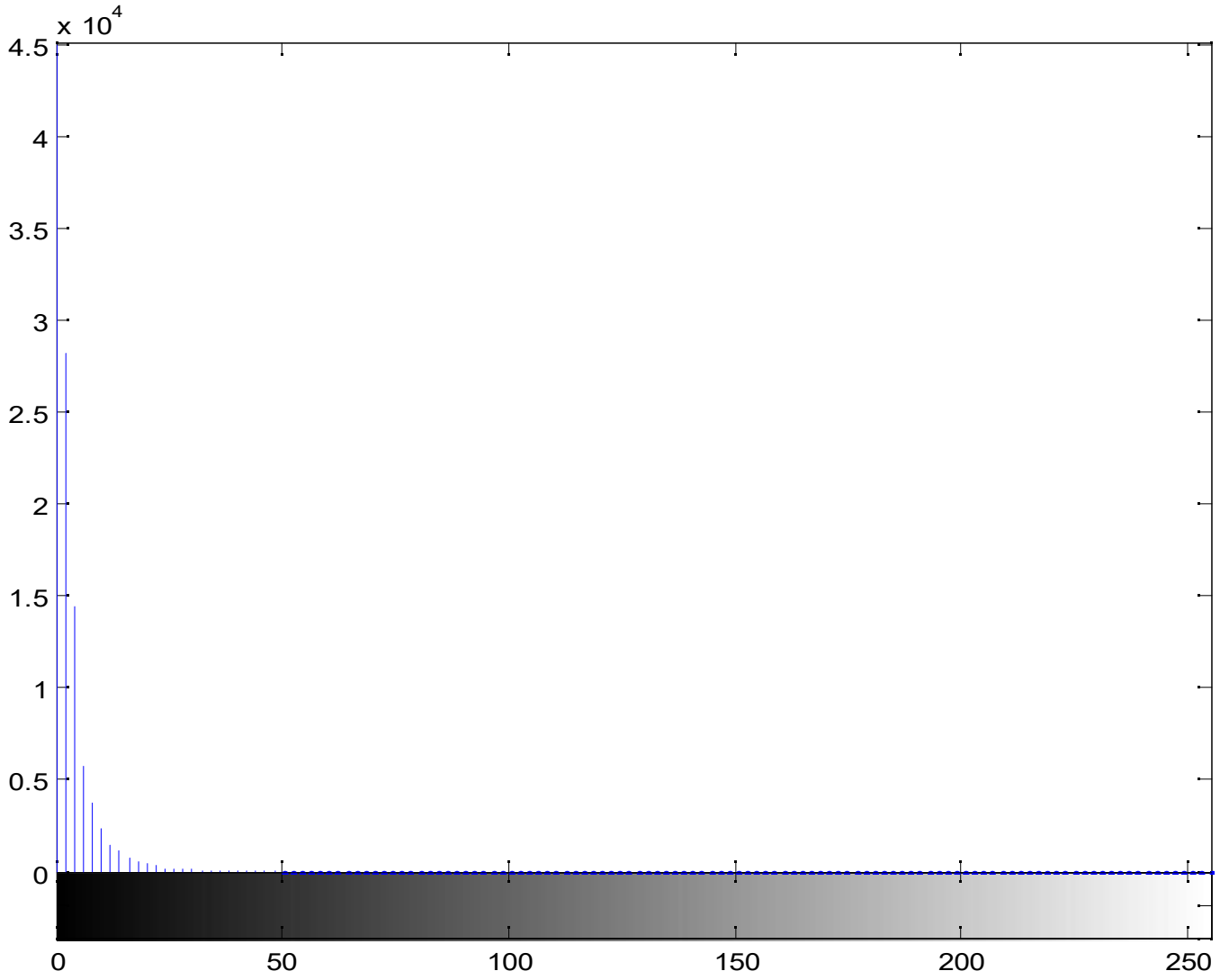
原理：对图像的一个像素的离散幅度的真实值，利用其相邻像素的相关性，预测它的下一个像素的可能值，再求两者差，对这种具有预测性质的差值，量化，编码，就可以达到压缩的目的.

例8 差分图像的直方图

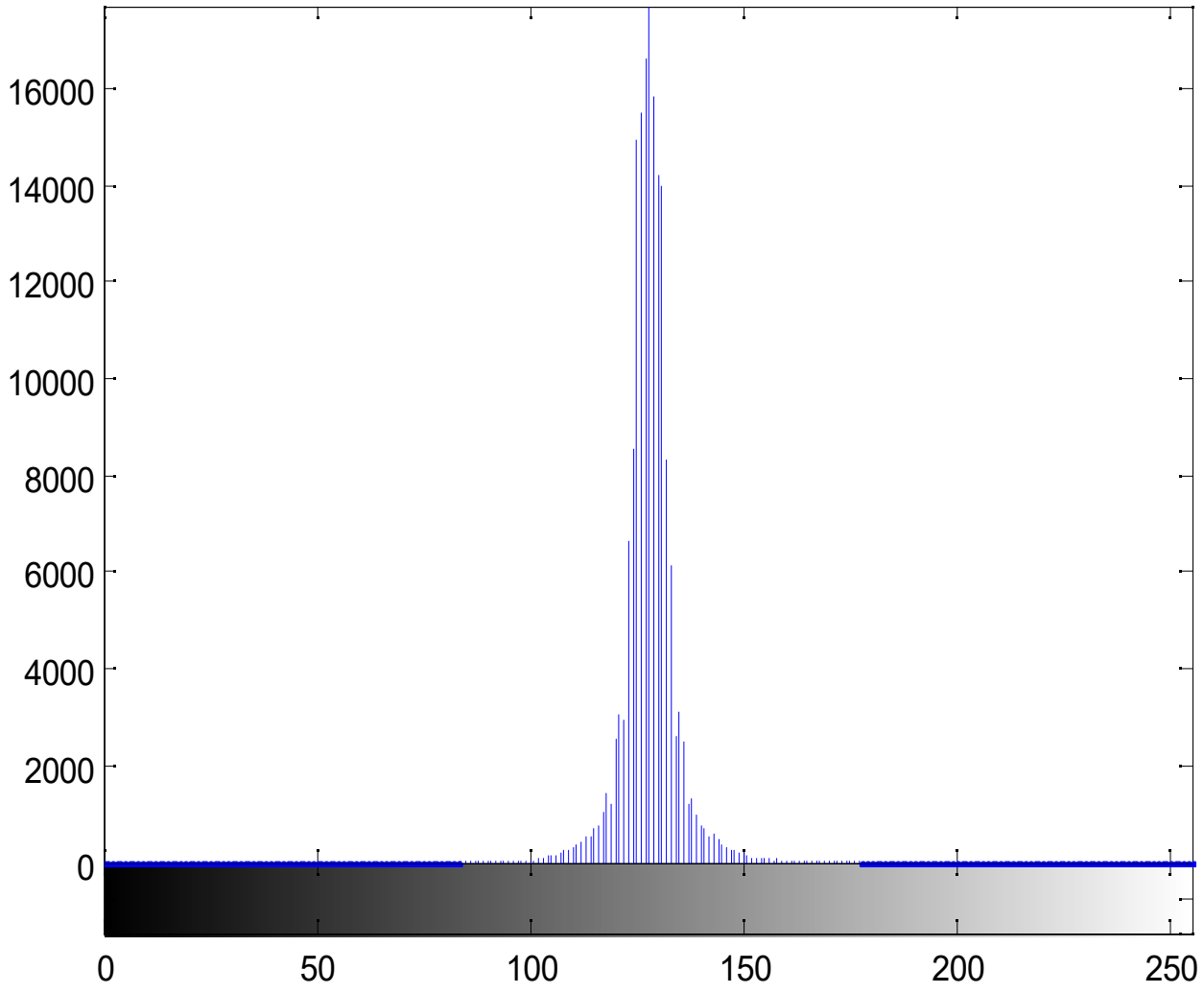
imhist(a);

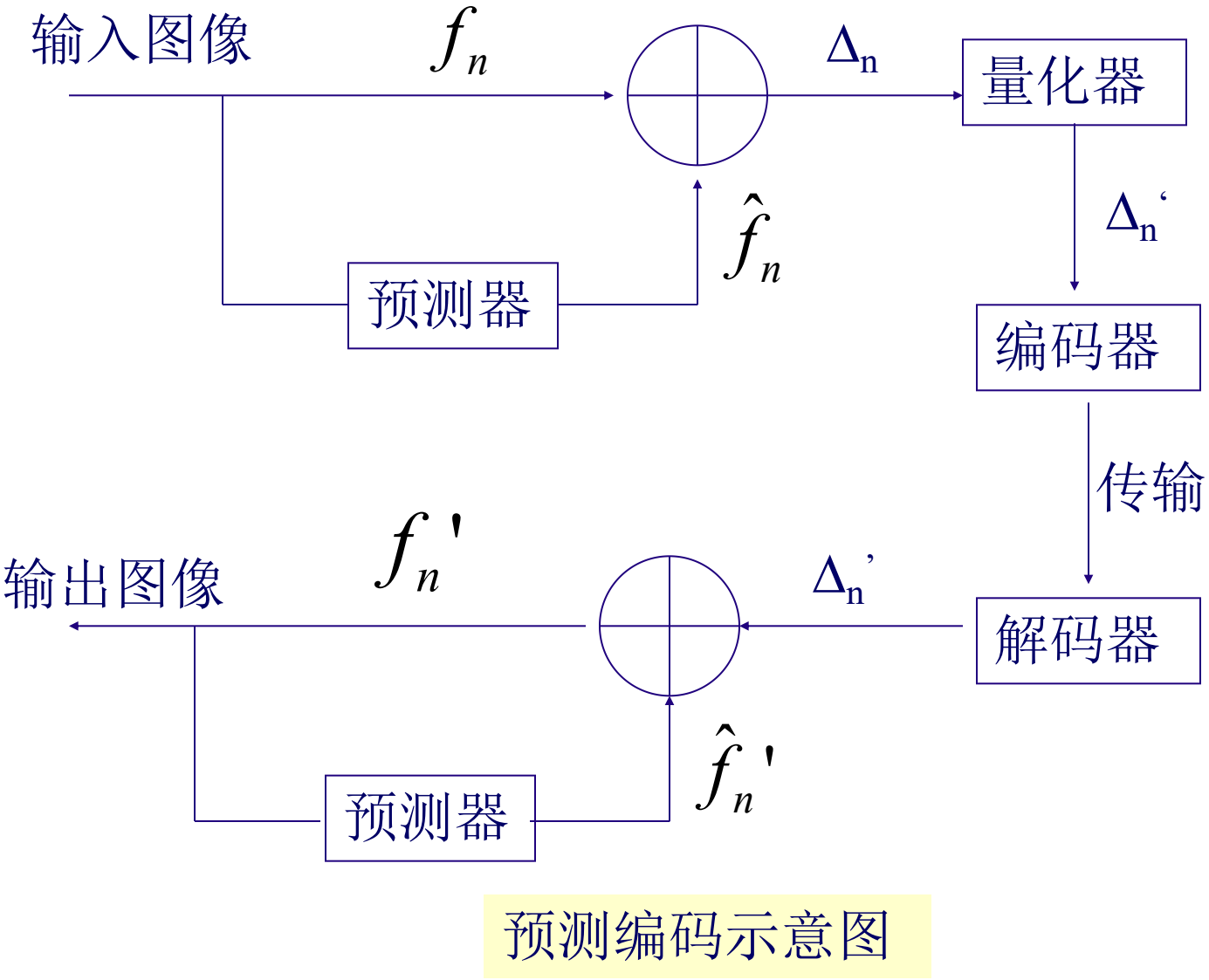


b=20*(a(i,j)-a(i+1,j+1));


x 10^4
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
0
50
100
150
200
250

```
a=imread('d.gif');
 for i=1:511
     for j=1:511
         e(i,j)=double(a(i,j))-double(a(i+1,j+1));
     end;
 end;
f=uint8(e+128);
imhist(f,256);
```


16000
14000
12000
10000
8000
6000
4000
2000
0
0
50
100
150
200
250


输入图像
f_n
Δ_n
量化器
Δ_n‘
编码器
传输
解码器
Δ_n’
f_n'
输出图像
$\hat{f}_n$
预测器
$\hat{f}_n$'
预测器

预测编码示意图

预测器：

$$\hat{f}_n = F(f_{n-1}, f_{n-2}, \cdots, f_{n-k})$$

$\hat{f}_n$ 是根据前面几个像素的亮度值 $f_{n-1}, f_{n-2}, \cdots, f_{n-k}$ 预测而得.

$$\Delta_n = f_n - \hat{f}_n$$

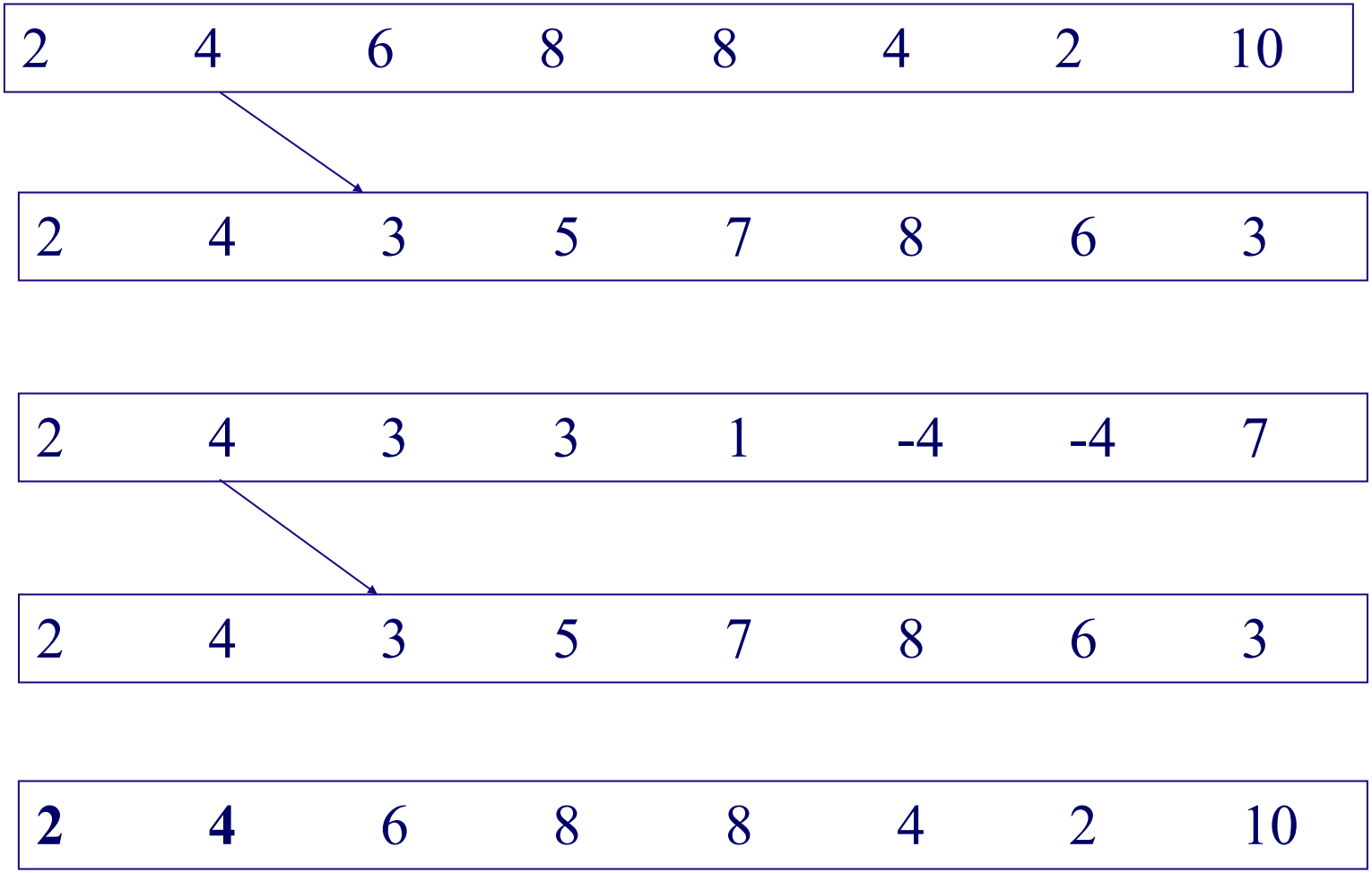
量化器： 对Δ_n进行舍入，整量化.

编码器： 可采用成熟的编码技术，如Huffman编码等.

解码器： 编码器的逆.

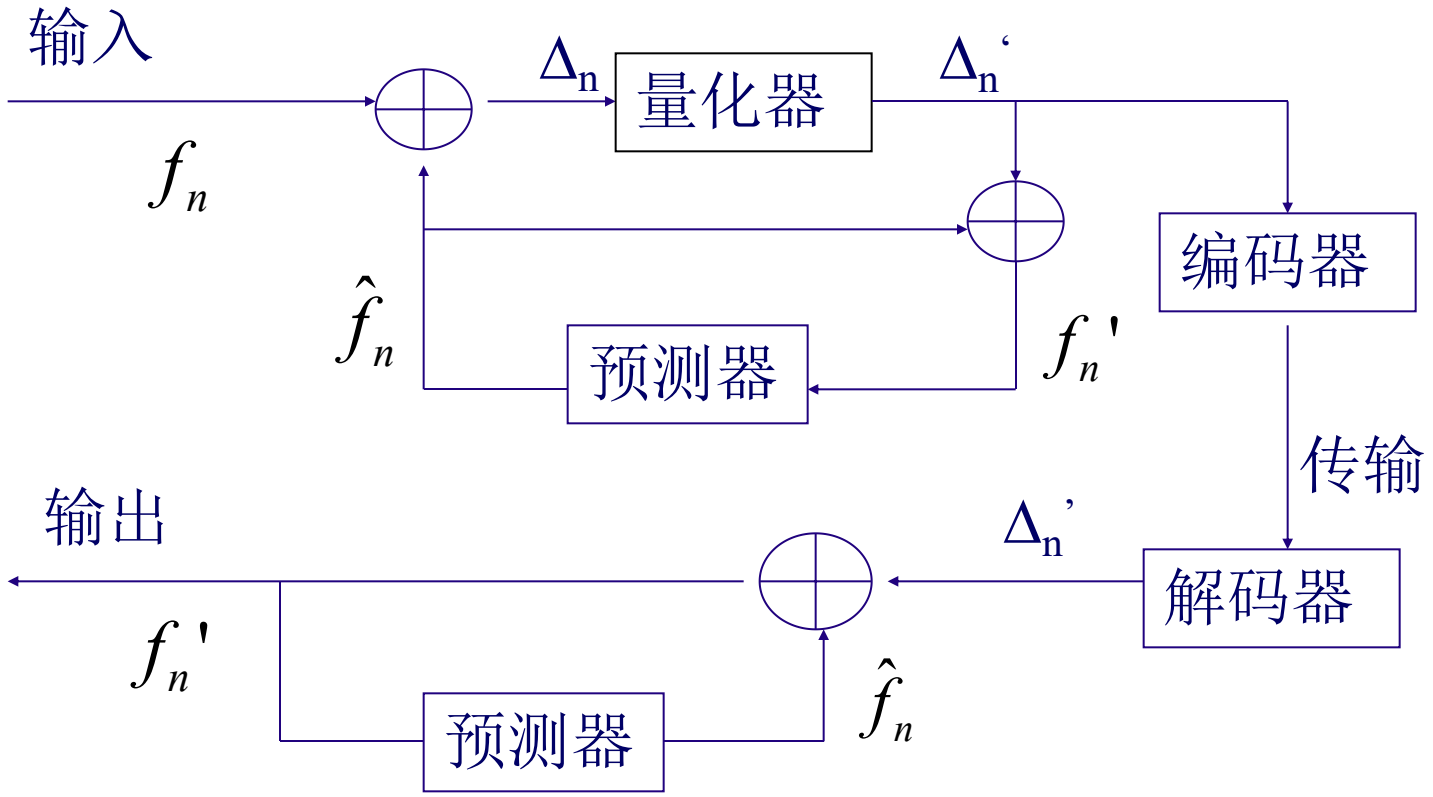
线性预测器： $\hat{f}_n = F(f_{n-1}, f_{n-2}, \cdots, f_{n-k}) = \sum_{l=n-k}^{n-1} a_l f_l, \sum_l a_l = 1$

例9



预测器 $\hat{f}_n = F(f_{n-1}, f_{n-2}) = \sum_{k=n-2}^{n-1} a_k f_k, \sum a_k = 0.5$

在预测编码中，最常用的是差分脉码调制（Differential Pulse Code Modulation, DPCM），原理图如下所示：

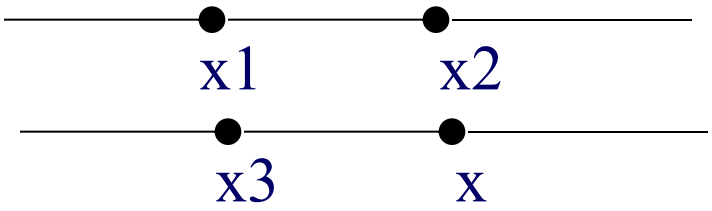


$$f_n' = \hat{f}_n + \Delta_n' \qquad \hat{f}_n = F(f'_{n-1}, f'_{n-2}, \cdots, f'_{n-k})$$

收端解码时的预测过程与发端相同，所用预测器也相同，收端输出的信号是发端的近似值，两者的误差是:

$$f_n' - f_n = \hat{f}_n + \Delta'_n - f_n = \Delta'_n - \Delta_n$$

1）多点预测



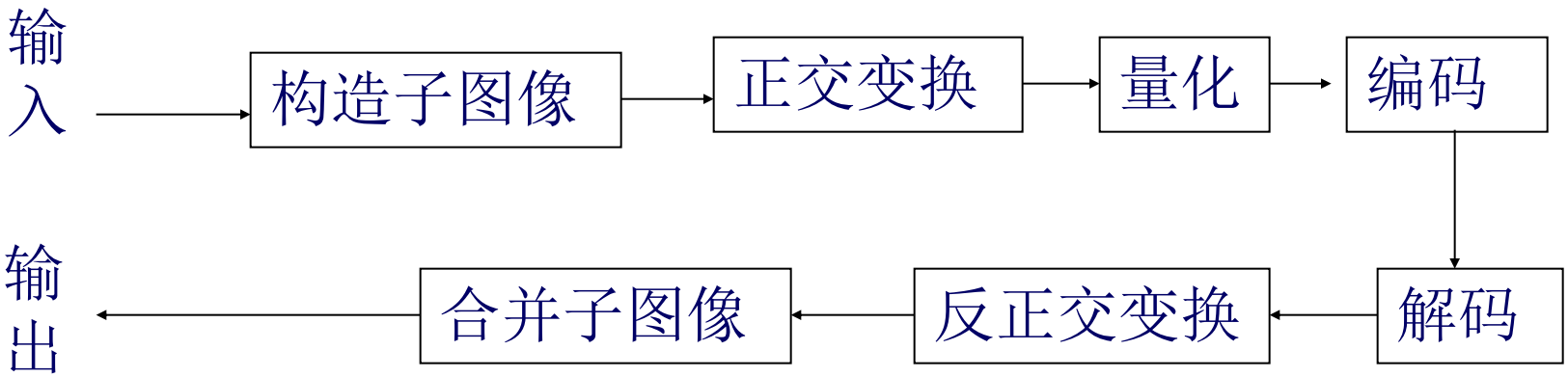
$$\hat{f}(x) = a_1 f(x1) + a_2 f(x2) + a_3 f(x3)$$

2)每行的最开始的几个像素无法预测，这些像素需要用其他方式编码，这是采用预测编码所需要的额外操作.

3）预测系数随着不同的图像而不同，但对每幅图像都计算预测系数太麻烦，也不现实，可参考前人得到的数据选择使用.在静止图像压缩的国际标准(JPEG)中，对这种方法的前置点形式以及预测系数有一推荐值可供参考.

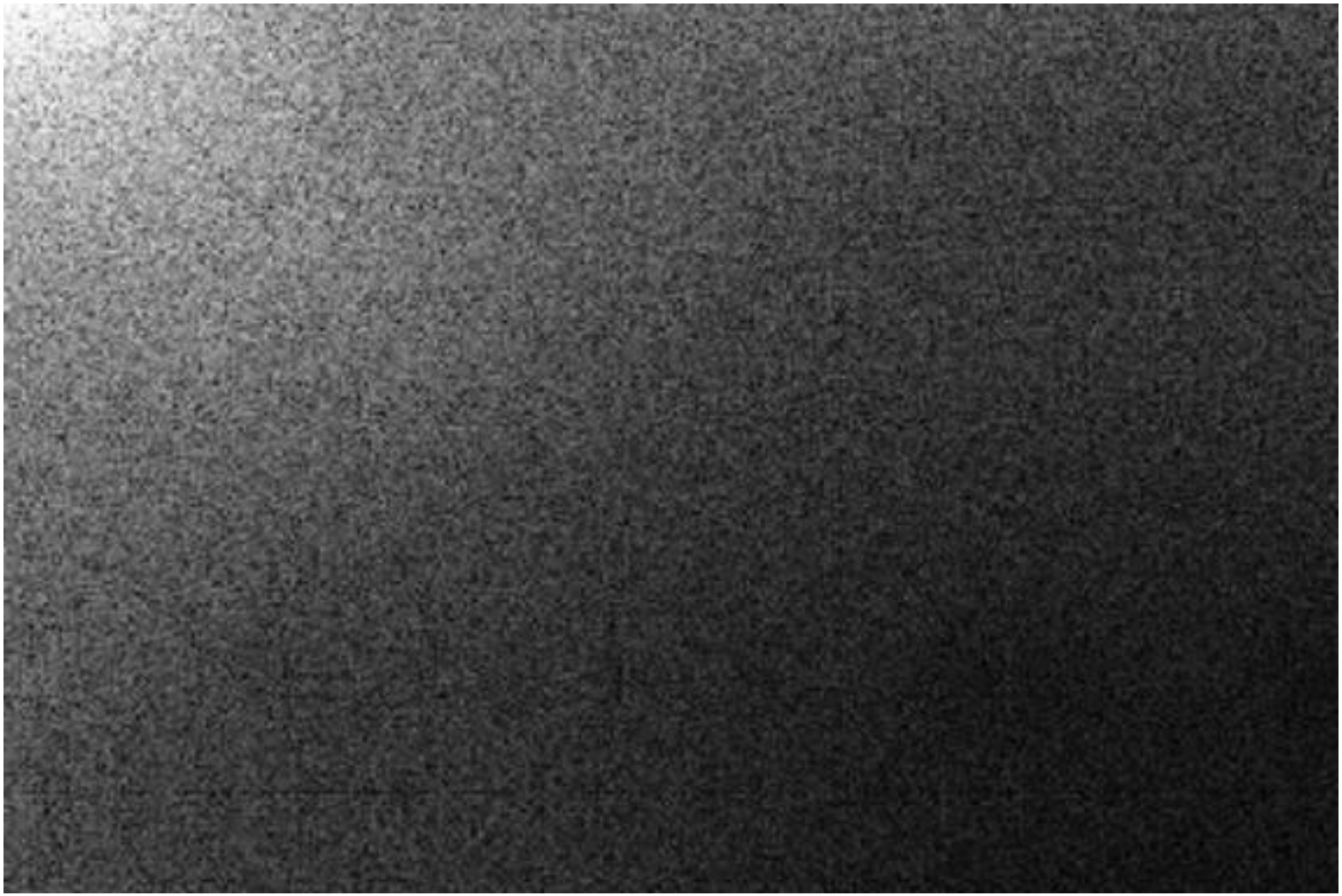
➢（块）变换编码

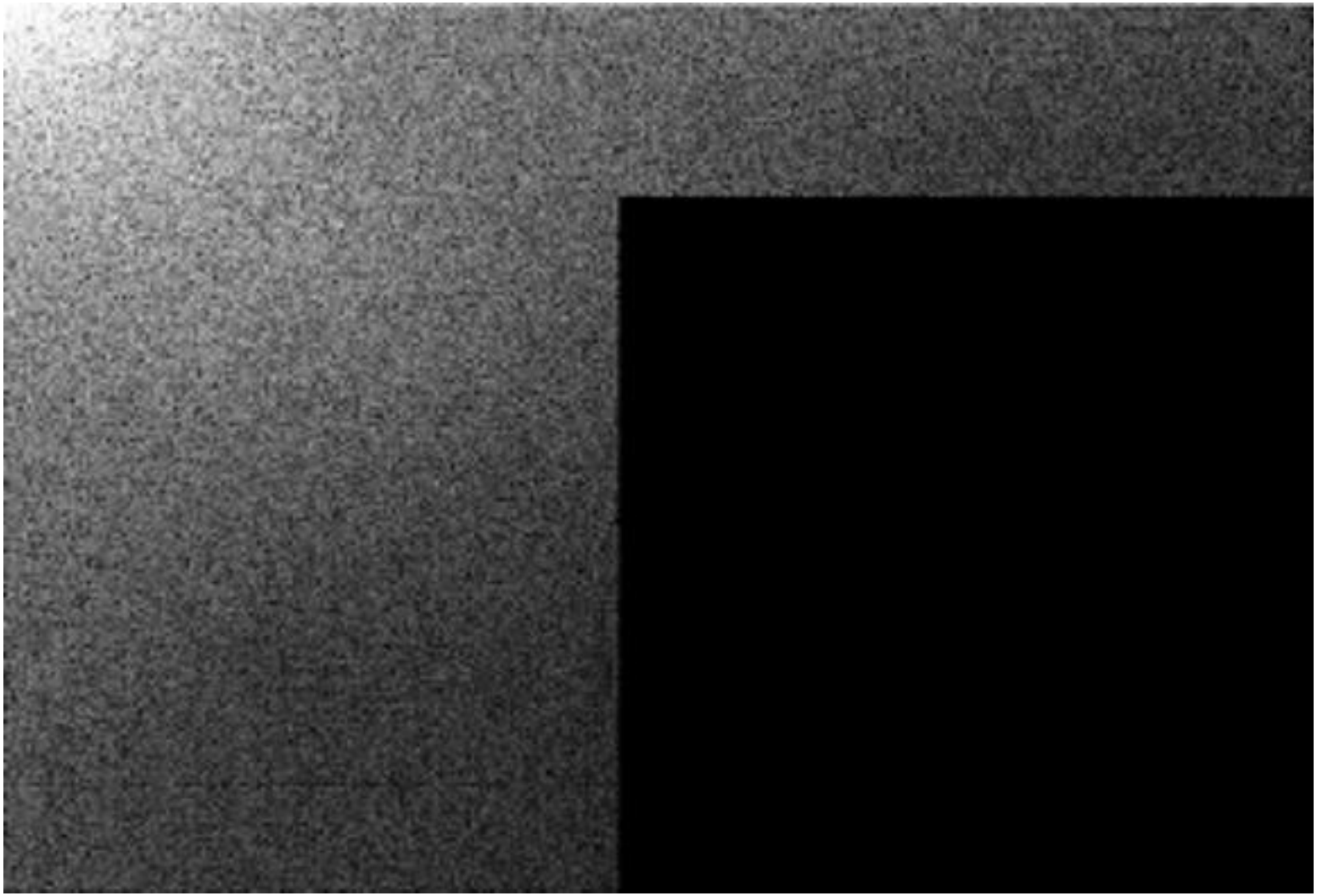
原理：图像数据经过正交变换后，其变换系数具有一定的相互独立性，（例如，对于FT来说，频普系数大的变换系数均集中在低频部分，而高频部分的幅值均很小，因而可以对低频的变换系数量化、编码和传输，对高频部分不处理，这样可以达到图像压缩的目的.

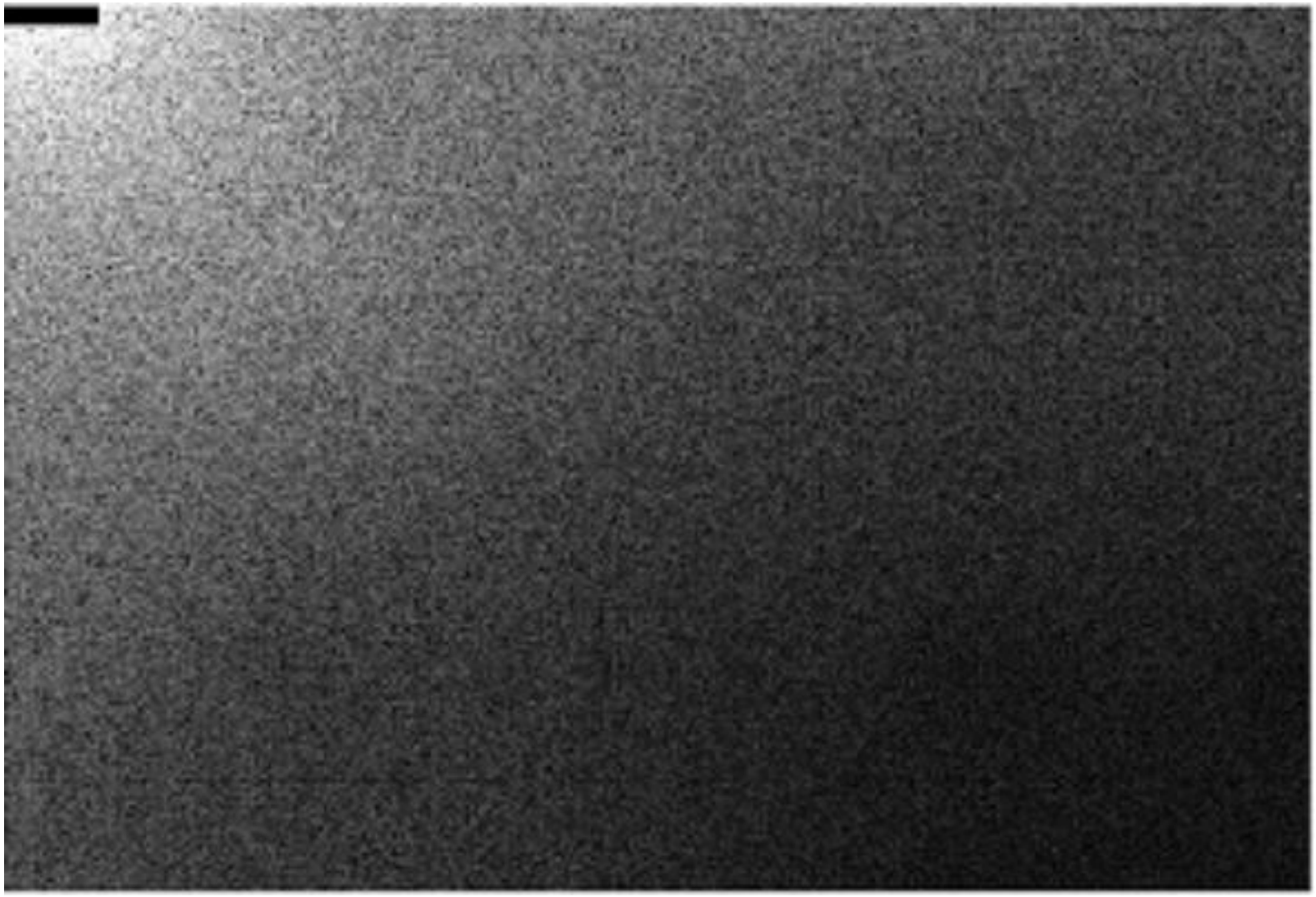


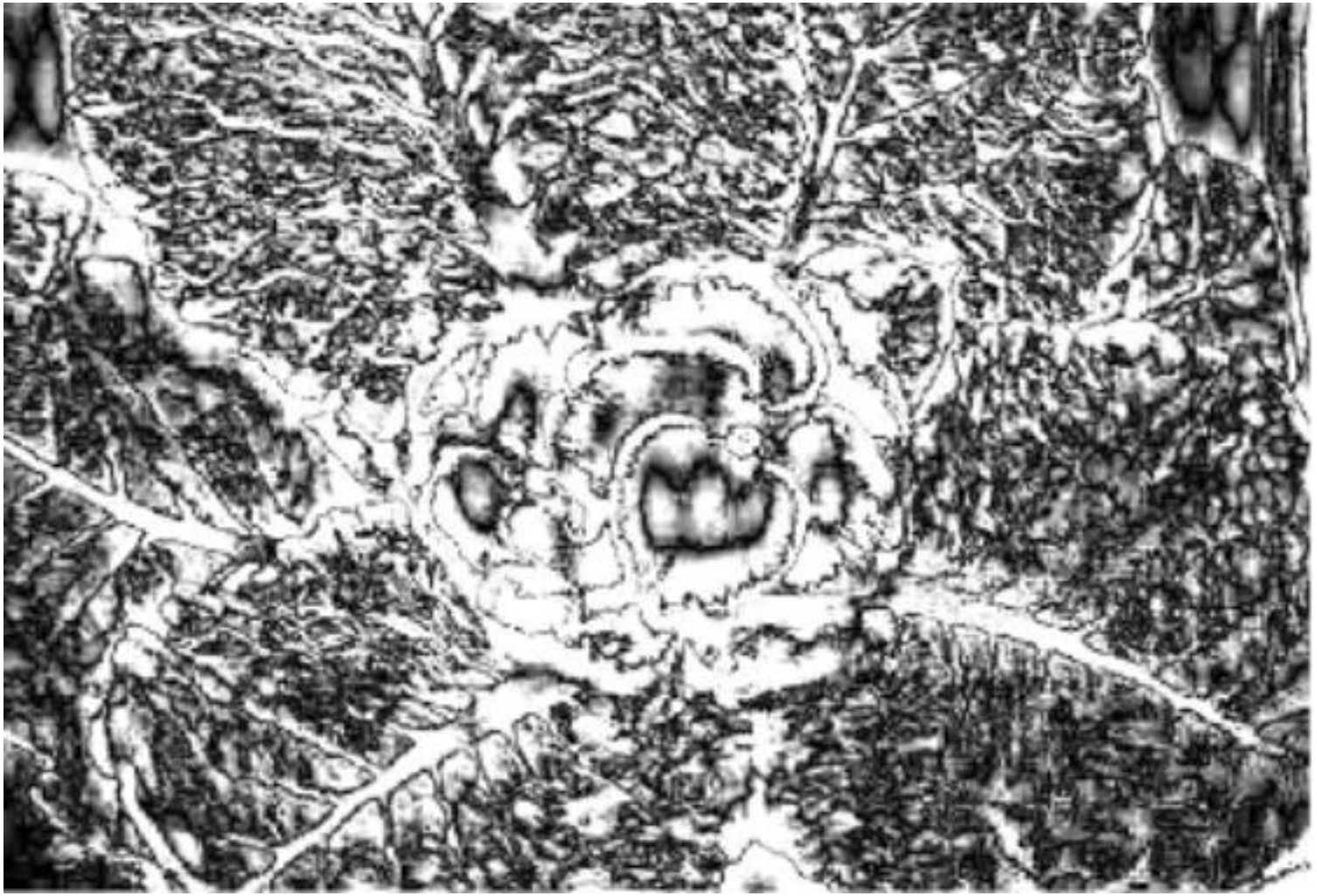
变换编码的一般系统框图

例10 (第5章，关于余弦变换的例子)









案例1 JPEG压缩

➢ 图像编码的国际标准

图像编码标准：

JBIG，H.26x，JPEG，MPEG

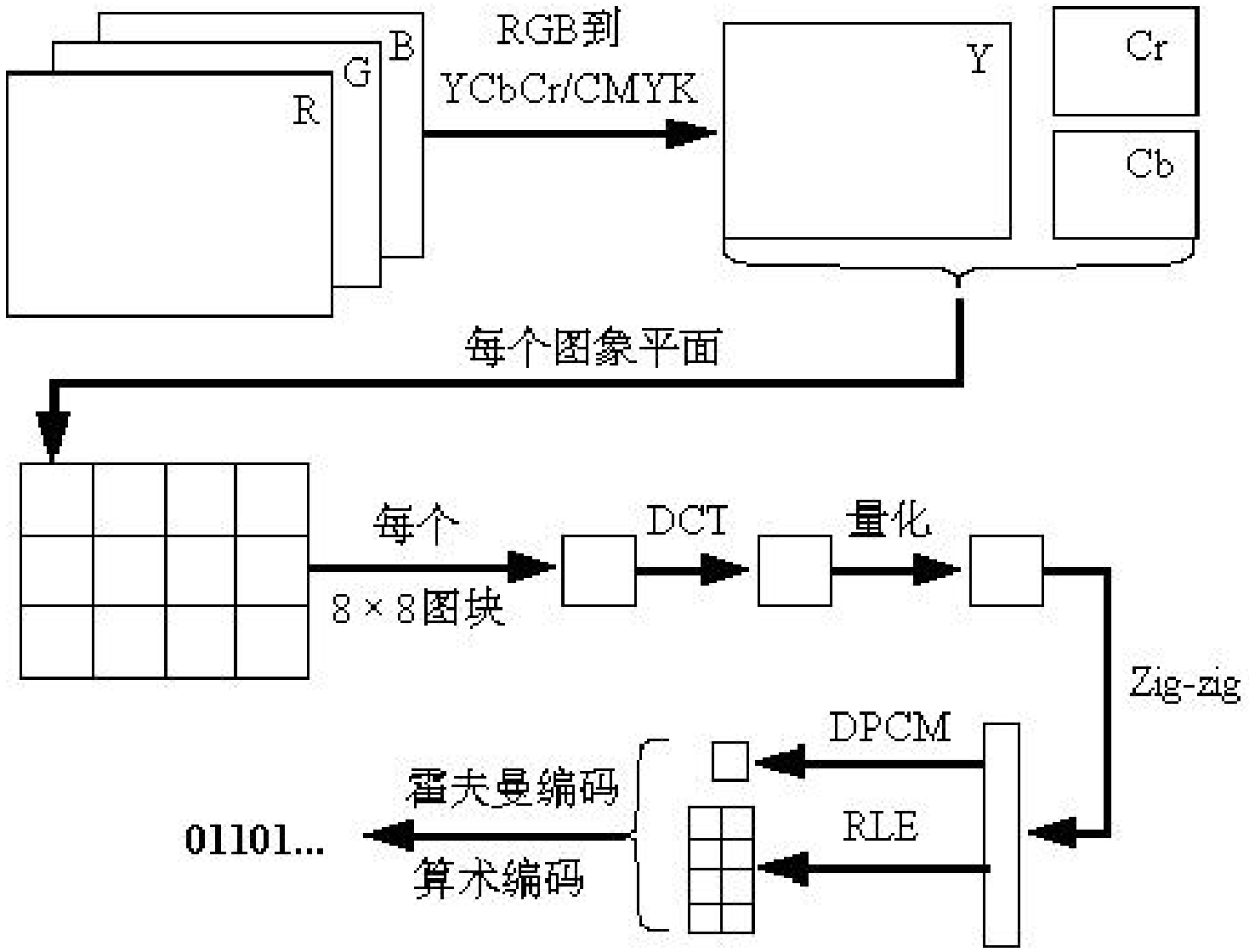
●JPEG

国际标准化组织（ID）和国际电报电话咨询委员会（CCITT）联合成立的专家组JPEG（Joint Photographic Experts Group）于1991年3月提出了ISO CDIO918号建议草案:多灰度静止图像的数字压缩编码（通常简称为JPEG标准）.这是一个适用于彩色和单色多灰度或连续色调静止数字图像的压缩标准.它包括基于DPCM（差分脉冲编码调制）、DCT（离散余弦变换）和Huffman编码的有损压缩算法两个部分.

JPEG压缩编码算法的主要计算步骤如下：

1. 正向离散余弦变换(FDCT).
2. 量化(quantization).
3. Z字形编码(zigzag scan).
4. 使用差分脉冲编码调制(differential pulse code modulation，DPCM)对直流系数(DC)进行编码.
5. 使用行程长度编码(run-length encoding，RLE)对交流系数(AC)进行编码.
6. 熵编码(entropy coding).

http://mti.xidian.edu.cn/multimedia/multi/

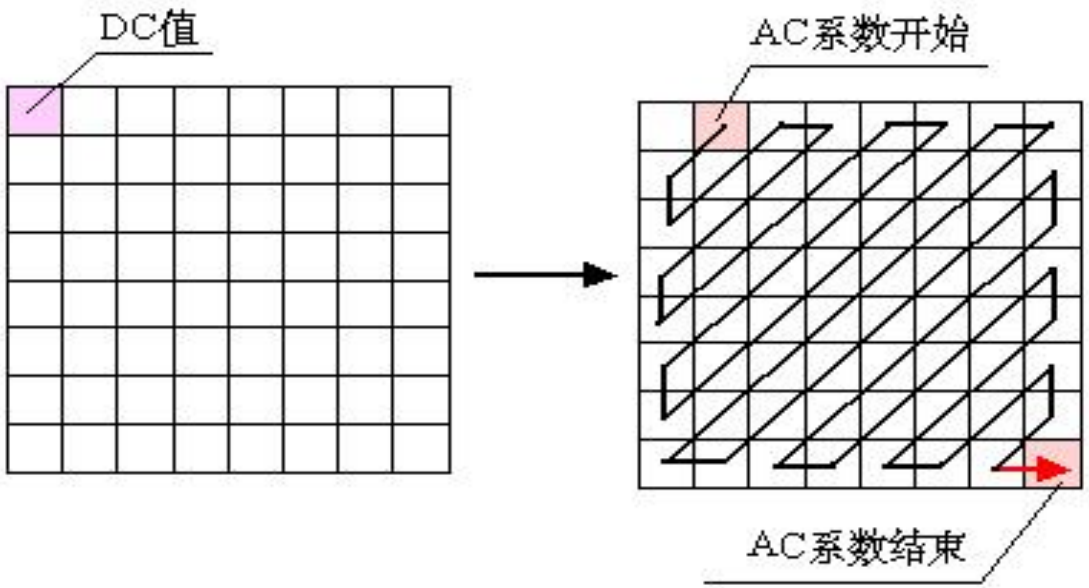

B
G
R
RGB到
YCbCr/CMYK
Y
Cr
Cb
每个图象平面
每个
8×8图块
DCT
量化
Zig-zig
DPCM
RLE
霍夫曼编码
算术编码
01101...

235.6	-1.0	-12.1	-5.20	2.1	-1.7	-2.7	1.3
-22.6	-18.5	-6.2	-3.2	-2.9	-0.1	0.4	-1.2
-10.9	-9.3	-1.6	1.5	0.2	-0.9	-0.6	-0.1
-7.1	-1.9	0.2	1.5	0.9	-0.1	0.0	0.3
-0.6	-0.8	1.5	1.6	-0.1	-0.7	0.6	1.3
1.8	-0.2	-1.6	-0.3	-0.8	1.5	1.0	-1.0
-1.3	-0.4	-0.3	-1.5	-0.5	1.7	1.1	-0.8
-2.6	1.6	-3.8	-1.8	1.9	1.2	-0.6	-0.4

FDCT系数

240	0	-10	0	0	0	0	0
-24	-12	0	0	0	0	0	0
-14	-13	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	0	0	0	0	0	0	0

逆量化后的系数



0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

量化DCT系数的序号

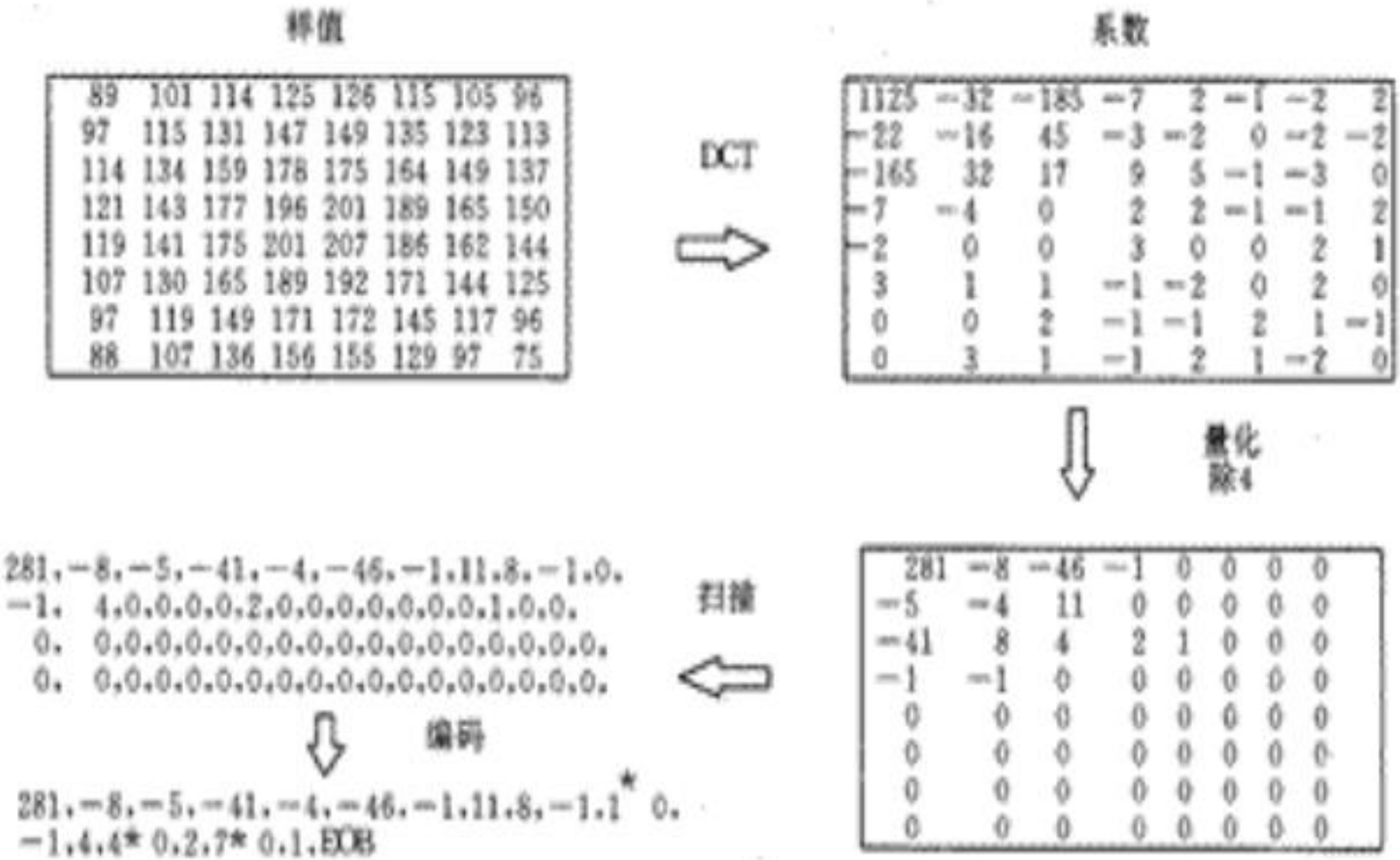


图 4-2 DCT 变换例

http://www.btc.sh.cn/wsxy/digi/d4z.htm

●MPEG

MPEG（Moving Pictures Experts Group）是 ISO/IEC/JTC/SC2/WG11的一个小组.它的工作兼顾了JPEG标准和CCITT专家组的H.261标准，于1990年形成了一个标准草案.

MPEG标准分成两个阶段:第一个阶段（**MPEG-I**）是针对传输速率为lMb/s到l.5Mb/s的普通电视质量的视频信号的压缩；第二个阶段（**MPEG-2**）目标则是对每秒30帧的720x572分辨率的视频信号进行压缩；在扩展模式下，MPEG-2可以对分辨率达1440Xl152高清晰度电视（HDTV）的信号进行压缩.

·MPEG-3：原本针对于HDTV(1920×1080)，后来被MPEG-2代替.
·MPEG-4：针对多媒体应用的图像编码标准.
·MPEG-7：基于内容表示的标准，应用于多媒体信息的搜索，过滤，组织和处理.

其他编码标准

- H.264

H.264是国际标准化组织（ISO）和国际电信联盟（ITU）共同提出的继MPEG-4之后的新一代数字视频压缩格式.

https://baike.baidu.com/item/H.264/1022230?fr=aladdin

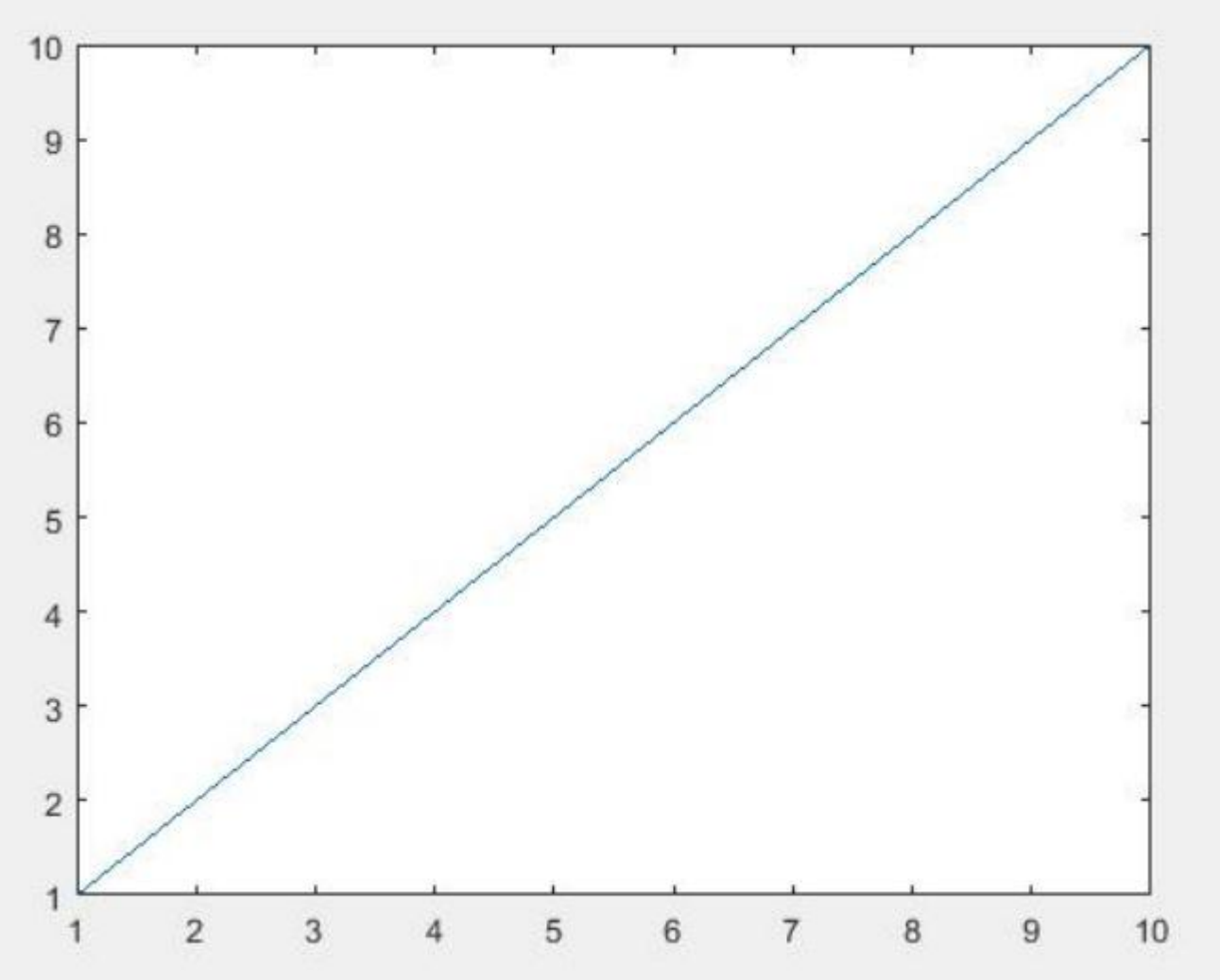
案例2 视频压缩

图像序列：具有次序关系的图像集合(时间次序、空间次序)。

视频：具有时间次序的、规格统一的图像序列(时间戳、视频帧)。

- 多帧图像读写

```
x=1:0.1:10;
filename='ding111.gif';
for n=1:0.5:5
      y = x.^n; plot(x,y);
      drawnow;frame = getframe(1);
      im = frame2im(frame); [A,map] = rgb2ind(im,256);
      if n == 1
            imwrite(A,map,filename,'gif','LoopCount',Inf,'DelayTime',1);
        else
            imwrite(A,map,filename,'gif','WriteMode','append','DelayTime',1);
      end
 end
```


10
9
8
7
6
5
4
3
2
1
1
2
3
4
5
6
7
8
9
10

imwrite

Write image to graphics file

Syntax

imwrite(A,filename) ;imwrite(A,map,filename);imwrite(___,fmt);
imwrite(___,Name,Value)

TIFF — Tagged Image File Format

'Compression' — Compression scheme
'packbits' | 'none' | 'lzw' | 'deflate' | 'jpeg' | 'ccitt' | 'fax3' | 'fax4'
Compression scheme, specified as the comma-separated pair consisting of 'Compression' and one of the following strings:

- 'packbits' (default for nonbinary images)
- 'none'
- 'lzw'
- 'deflate'
- 'jpeg'
- 'ccitt' (binary images only, and the default for such images)
- 'fax3' (binary images only)
- 'fax4' (binary images only)
- 'jpeg' is a lossy compression scheme; other compression modes are lossless. Also, if you specify 'jpeg' compression, you must specify the 'RowsPerStrip' parameter and the value must be a multiple of 8.

Example: 'Compression','none'

- Matlab图像序列和电影

  自学

  im2frame,frame2im

  tifs2seq,seq2tifs

  tifs2movie,movie2tifs

  tifs2cv,cv2tifs (Page197)

  movie2avi

  implay

- 时间冗余和运动补偿

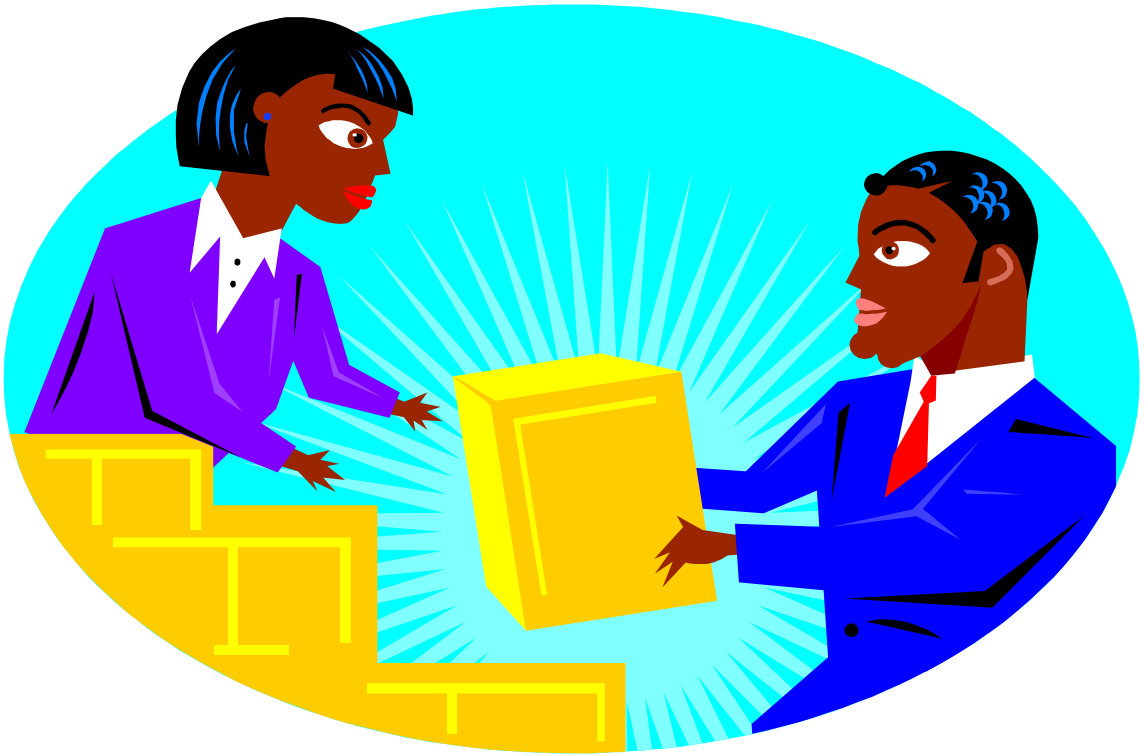
  自学

  I帧：无预测压缩的帧称为内帧或独立帧。关键帧, I-frame。

  P帧：基于前一帧编码的帧称为预测帧。P-frame。

  B帧：基于前驱帧和后继帧进行预测的帧称为双向帧。B-frame。

回答问题



作业：

8.1, 8.5, 8.10, 8.30, 8.34.

注：红色标注的题目请尝试完成

本次(第四次)作业截止日期:2021年11月7日24:00点