



数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

1.2 数字图像的存储与读写

1.3 数字图像处理系统简介

1.4 数字图像处理应用

演示文稿说明：

本讲内容以板书为主，*ppt* 演示为辅。





第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

- 图： 物体反射光 \ 透射光 \ 发光物体本身发射光（能量）的分布（图像场），是客观存在；
- 像： 人的视觉系统所接受的图在人脑中所形成的印象或认识，是人的主观感觉；
- 图像：是二维或三维景物呈现在人心目中的影像。图和像的有机结合，既反映物体的客观存在，又体现人的心理因素，是对客观存在的物体的一种相似性的生动模仿或描述，是对物体的一种不完全、不精确，但在某种意义上是适当的表示。



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

图像是与之对应景物的一个表示，包含了描述其所代表的景物的信息。一幅三维空间的、动态的图像表示为：

$$I = f(x, y, z, t)$$

I 是物体辐射能量， x 、 y 、 z 是空间变量， t 是时间变量。

在实际应用中，图像可以借助不同的方式来获得：

- 一定波谱段内的电磁能量成像：在低频端，如红外线、微波等；在高频端，如紫外光、X光、 γ 射线、宇宙射线等。此时图像表示为：

$$I = f(x, y, z, \lambda, t)$$

λ 是波长。 λ 不同，物体的反射、发射或吸收特性不同。当对可见光成像时，对应客观景物被观察到的辐射能量即为亮度。一般情况下只考虑平面、单色、静止图像，图像表示为一个二维函数：

$$I = f(x, y)$$

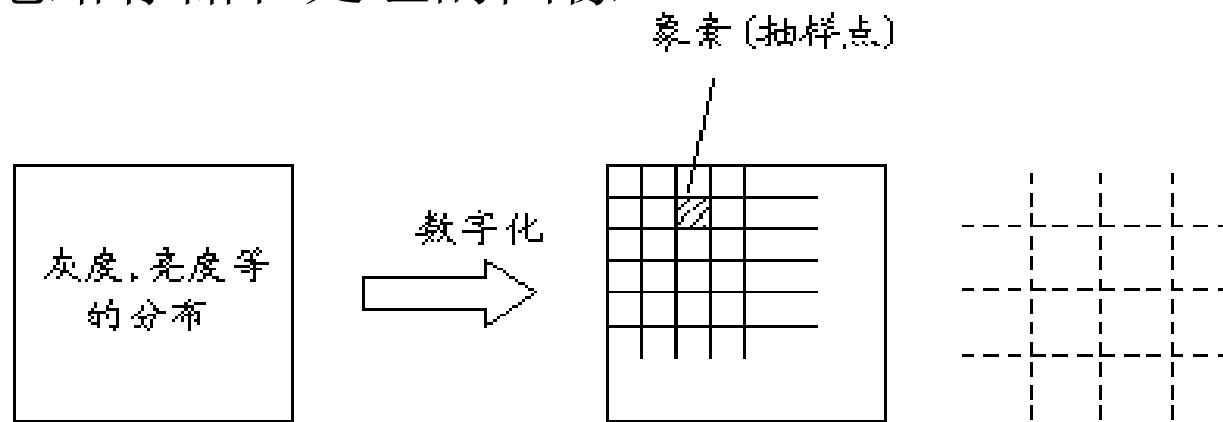
- 声波、超声波、电子束、磁共振成像等。



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

模拟图像与数字图像：数字图像是由模拟图象数字化得到的，以像素为基本元素、可以用数字计算机或数字电路存储和处理的图像。



把一幅图画分割成如图所示的一个个小区域（像素），并将各小区域灰度用整数来表示，形成一幅点阵式的数字图像。它包括采样和量化两个过程。像素的位置和灰度就是像素的属性。





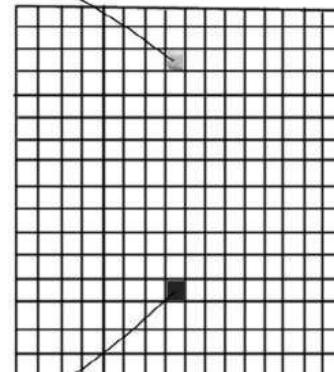
第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

一幅模拟图像的坐标及幅度都是连续的，为了把它转换为数字形式，必须对坐标和幅度都作离散化操作。数字化坐标值称为采样，它确定了图像的空间分辨率；数字化幅度值称为量化，它确定了图像的幅度分辨率。



物理图象



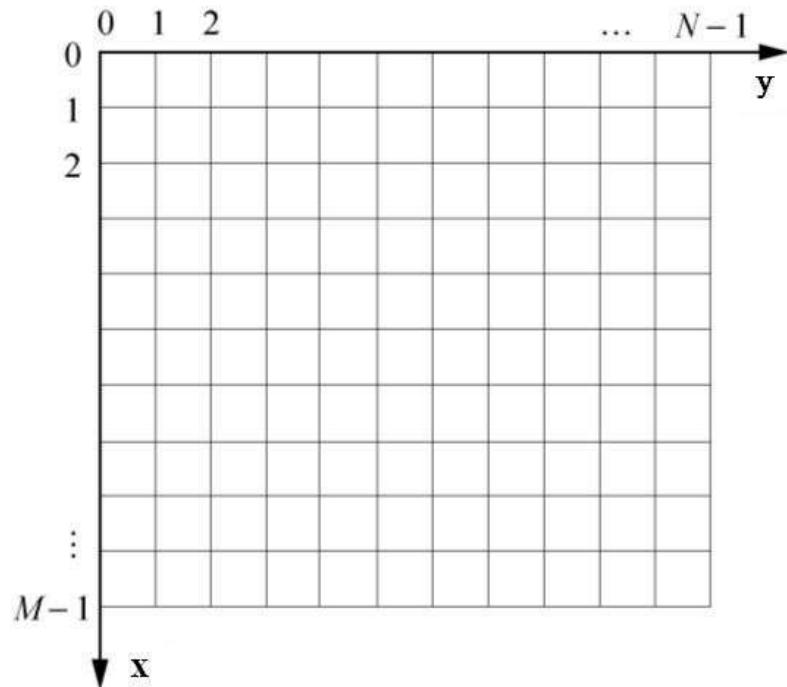
数字图象



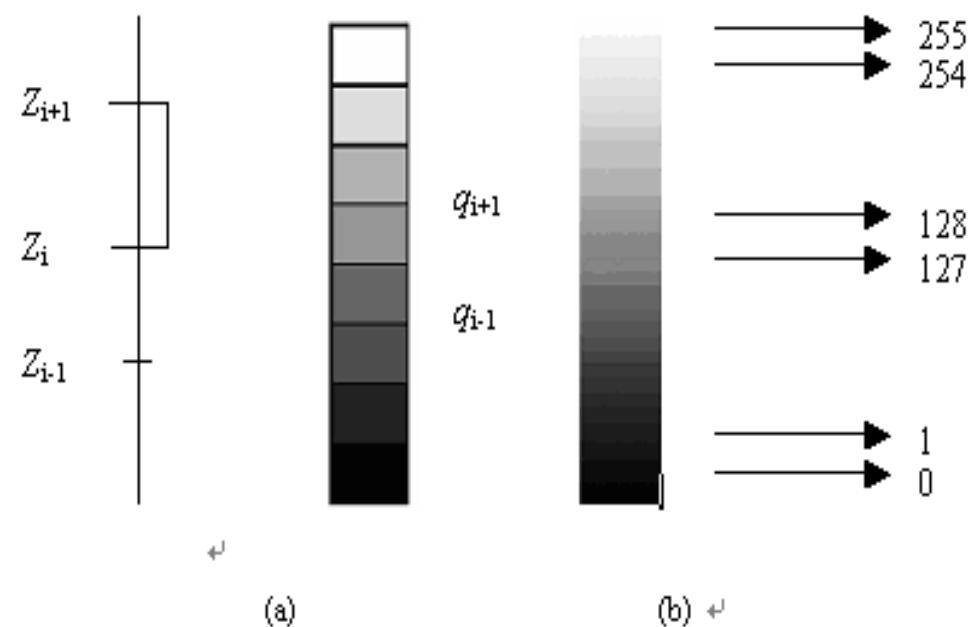


第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示



图像采样过程示意图



(a) 图像量化过程示意图 (b)8bit 量化



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

对于灰度图像，量化是对采样所得的离散样本点上的灰度值进行离散化，将原图像的连续灰度用 $L = 2^k$ (k 为整数)个等间距的灰度级进行表示。连续图像被取样与量化后可以用一个矩阵来表示，即

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}$$

$f(x,y)$ 被称为数字图像，矩阵中的每一个元素称为像素(Pixel)。

若一幅数字图像的量化灰度级数 $L=256$ 级($k=8$)，灰度取值范围是0~255的整数，由于用8bit就能表示灰度图像像素的灰度值，因此常称8 bit 量化。一幅大小为 $M \times N$ 、灰度级数为 L 的图像所需的存储空间，即图像的数据量，大小为 $b = M \times N \times k$ (bit)



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示



(a)



(b)

18	17	19	17	21	29	45	59	65	59	58	66	67	61	69	60
22	20	20	17	19	25	51	65	82	90	84	74	73	78	57	56
27	23	23	18	17	21	42	47	66	90	97	90	84	86	58	61
28	25	24	21	19	21	24	24	30	50	77	95	93	84	79	77
26	24	24	23	22	23	26	38	37	28	43	77	93	88	102	91
24	20	20	21	22	23	40	68	75	47	29	48	80	97	109	97
23	16	15	17	19	19	36	55	73	68	44	33	58	92	108	103
23	14	11	13	15	15	16	12	36	69	64	35	42	77	108	110
18	21	20	19	16	7	8	14	31	60	63	30	32	79	106	118
19	18	13	13	18	17	5	11	23	48	57	38	45	84	122	128
21	18	10	13	28	35	29	42	51	53	46	40	63	104	140	137
22	24	15	18	35	46	58	77	82	60	35	42	90	140	152	140
21	27	19	21	35	44	46	53	52	38	36	72	131	172	164	146
20	26	24	31	46	54	28	14	13	31	70	128	174	187	180	156
20	26	36	60	88	101	74	55	63	99	138	178	196	186	190	163
22	28	50	91	133	152	149	140	160	189	197	201	198	182	192	165

(c)

数字图像表示实例

- (a) 256 级灰度图象 (b) 子图 (c) 子图对应的量化数据



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

空间和幅度分辨率同时变化所产生的效果

图像幅度分辨率

图像空间分辨率



(a)

(b)

(c)



(d)

(e)

(f)

图 2.4.3 图像幅度分辨率变化所产生的效果

图 2.4.2 图像空间分辨率变化所产生的效果



图 2.4.4 图像空间和幅度分辨率同时变化所产生的效果





第1章 数字图像处理的基本知识

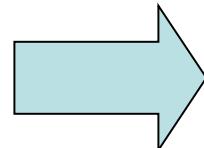
1.1 图像的概念、数字图像及表示

灰度图像：当一幅图像有 2^k 灰度级时，通常称该图像是 k 比特图像

。

例如，一幅图像有 256 个可能的灰度级，称其为 8 比特图像。灰度图像矩阵元素的取值范围通常为 $[0, 255]$ ，因此其数据类型一般为 8 位无符号整数，这就是人们经常提到的 256 级灰度图像。“0”表示纯黑色，“255”表示纯白色，中间的数字从小到大表示由黑到白的过渡色

。



$$I = \begin{bmatrix} 0 & 150 & 200 \\ 120 & 50 & 180 \\ 250 & 220 & 100 \end{bmatrix}$$

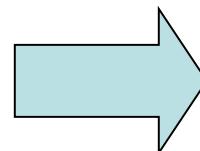
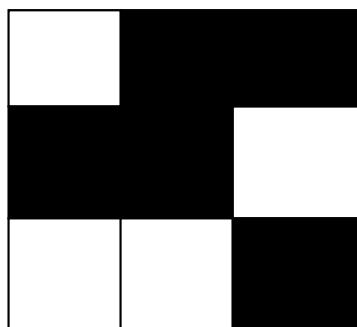




第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

二值图像：灰度值只由 0、1 两个值构成，“0”代表黑色，“1”代表白色。由于每一像素的取值仅有 0、1 两种可能，所以计算机中二值图像的数据类型通常为一个二进制位。二值图像通常用于文字、线条图的扫描识别(OCR)和掩模图像的存储。二值图像可以看成是灰度图像的一个特例。



$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

RGB 图像用来表示彩色图像。它分别用红 (R)、绿 (G)、蓝 (B) 三基色的组合来表示每个像素的颜色。图像中每一个像素的颜色值 (由 RGB 三基色表示) 直接存放在图像矩阵中，由于每一像素的颜色需由 R、G、B 三个分量来表示，因此 RGB 图像的图像矩阵与其他类型的图像矩阵不同，是一个三维矩阵，可用 $M \times N \times 3$ 表示， M 、 N 分别表示图像的行、列数，三个 $M \times N$ 的二维矩阵分别表示各个像素的 R、G、B 三个颜色分量。每个颜色分量的数据类型一般为 8 位无符号整型。

The diagram illustrates the conversion of a 3x3 pixel image into its corresponding RGB matrix. On the left, a 3x3 grid of colored squares represents the image. The colors are: Row 1: Red, Brown, Magenta; Row 2: Yellow, Green, Blue; Row 3: Magenta, Cyan, Blue. A large blue arrow points from this image to the right, where three matrices are defined:

$$R = \begin{bmatrix} 255 & 240 & 240 \\ 255 & 0 & 80 \\ 255 & 0 & 0 \end{bmatrix} \quad G = \begin{bmatrix} 0 & 160 & 80 \\ 255 & 255 & 160 \\ 0 & 255 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 80 & 160 \\ 0 & 0 & 240 \\ 255 & 255 & 255 \end{bmatrix}$$



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示



176 177 176 175 177 174 175 172 176 172 167 165 164 182 206 221 224 229
177 177 175 173 178 175 170 177 178 172 167 166 164 180 203 217 224 231
173 175 174 173 179 182 177 175 176 174 167 160 161 178 201 216 222 227
172 175 177 177 178 173 177 175 176 168 163 161 163 182 207 221 224 229
174 174 173 179 174 174 170 171 175 168 162 157 160 183 206 219 223 230
173 176 173 170 173 174 170 171 169 168 162 165 156 179 204 219 223 232
168 170 169 163 170 169 163 163 166 164 154 151 155 173 198 215 222 231
169 169 173 167 167 169 166 167 162 160 158 147 149 169 199 217 220 228
170 172 169 165 167 166 164 164 162 152 154 146 149 170 199 217 220 230
171 171 169 167 170 169 167 165 160 154 151 144 148 171 199 214 220 230

灰度图像：单色，一元组



(207, 137, 130)	(220, 179, 163)	(215, 169, 161)	(210, 179, 172)	(210, 179, 172)
(207, 154, 146)	(217, 124, 121)	(226, 144, 133)	(226, 144, 133)	(224, 137, 124)
(227, 151, 136)	(227, 151, 136)	(226, 159, 142)	(227, 151, 136)	(230, 170, 154)
(231, 178, 163)	(231, 178, 163)	(231, 178, 163)	(236, 187, 171)	(236, 187, 171)
(239, 195, 176)	(239, 195, 176)	(240, 205, 187)	(239, 195, 176)	(231, 138, 123)
(217, 124, 121)	(215, 169, 161)	(216, 179, 170)	(216, 179, 170)	(207, 137, 120)
(159, 51, 71)	(189, 89, 101)	(216, 111, 110)	(217, 124, 121)	(227, 151, 136)
(227, 151, 136)	(226, 159, 142)	(226, 159, 142)	(237, 159, 135)	(237, 159, 135)
(231, 178, 163)	(236, 187, 171)	(231, 178, 163)	(236, 187, 171)	(236, 187, 171)
(236, 187, 171)	(239, 195, 176)	(239, 195, 176)	(236, 187, 171)	(227, 133, 118)
(213, 142, 135)	(216, 179, 170)	(221, 184, 170)	(190, 89, 89)	(204, 109, 113)
(204, 115, 118)	(189, 85, 97)	(159, 60, 78)	(136, 38, 65)	(160, 56, 75)

彩色图像：RGB 三色，三元组





第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

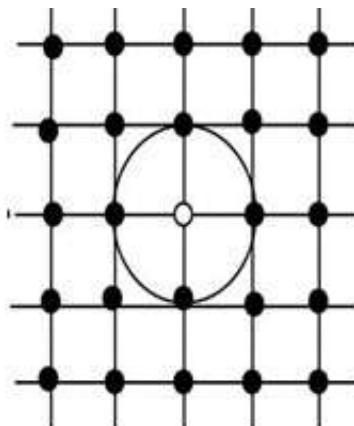
一幅图像由基本单元像素组成，像素间存在着一定的联系，包括像素的邻域，邻接和连通，以及像素间的距离。

➤ 像素的邻域

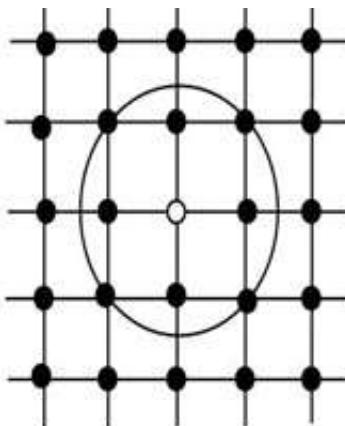
4—邻域 $N_4(p)$: 包括水平和垂直4个邻近像素。

8—邻域 $N_8(p)$: 包括水平、垂直和对角的8个邻近像素。

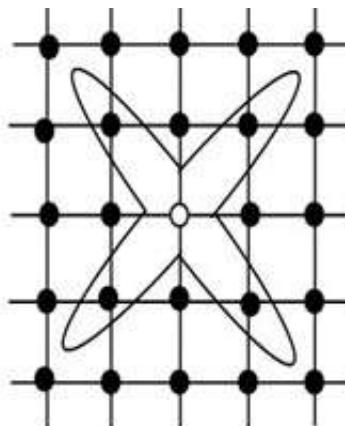
对角邻域 $N_D(p)$: 包括对角的4个邻近像素。



(a)4-邻域



(b)8-邻域



(c)对角邻域





数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第 1 章 数字图像处理的基本知识

第 2 讲 像素间的基本关系、数字图像文件的存储格式

演示文稿说明：

- 本讲内容以板书为主，*ppt* 演示为辅；



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

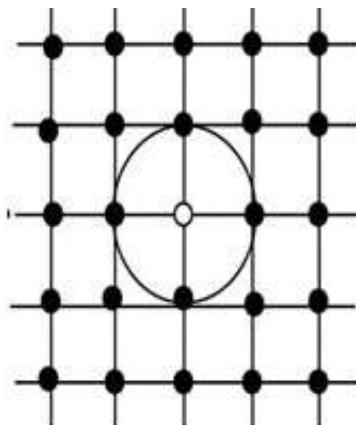
一幅图像由基本单元像素组成，像素间存在着一定的联系，包括像素的邻域，邻接和连通，以及像素间的距离。

➤ 像素的邻域

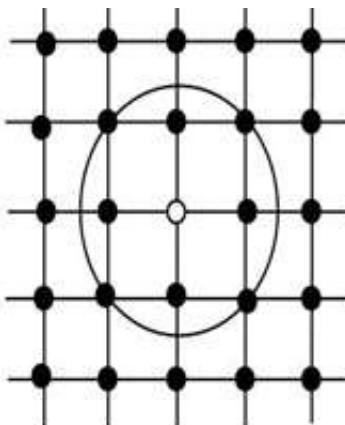
4—邻域 $N_4(p)$: 包括水平和垂直4个邻近像素。

8—邻域 $N_8(p)$: 包括水平、垂直和对角的8个邻近像素。

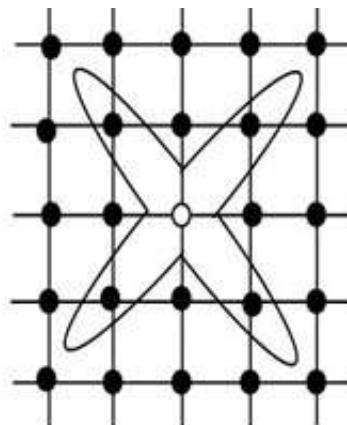
对角邻域 $N_D(p)$: 包括对角的4个邻近像素。



(a)4-邻域



(b)8-邻域



(c)对角邻域





第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

➤ 邻接与连通

像素间邻接的两个必要条件：

- ① 一个像素在另一个像素的邻域中；
- ② 两个像素的灰度值满足特定的相似准则。

邻接类型：

- ① 4- 邻接
- ② 8- 邻接
- ③ m- 邻接



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

令 V 表示关于邻接的灰度值集合（相似准则）。

① 4- 邻接

像素 p 和 q 的灰度值都属于 V ，且 q 在集合 $N_4(p)$ 中，则称像素 p 和 q 是 4- 邻接的。

② 8- 邻接

像素 p 和 q 的灰度值都属于 V ，且 q 在集合 $N_8(p)$ 中，则称像素 p 和 q 是 8- 邻接的。

③ m- 邻接

像素 p 和 q 的灰度值都属于 V ，若满足下列条件之一：

- q 在集合 $N_4(p)$ 中；
- q 在集合 $N_D(p)$ 中，且集合 $N_4(p) \cap N_4(q)$ 中的像素的灰度值不属于 V 。则称像素 p 和 q 是 m- 邻接的。



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

0 1 1

0 1 0

0 0 1

0 1 -1

0 1 0

0 0 1

0 1 -1

0 1 0

0 0 1

0 1 -1

0 1 0

0 0 1

(a) 像素排列
邻接

(b) 4- 邻接

(c) 8- 邻接

(d) m -
邻接



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

介绍像素间的连通之前，先定义像素间的通路：

从坐标为 (x_0, y_0) 的像素 p 到坐标为 (x_n, y_n) 的像素 q 的一条通路由像素序列组成，对应坐标序列为

$$(x_0, y_0), (x_1, y_1), \dots, (x_i, y_i), \dots, (x_n, y_n)$$

这里， (x_i, y_i) 与 (x_{i-1}, y_{i-1}) 相邻接，其中 $1 \leq i \leq n$ ， n 为通路长度。若 $(x_0, y_0) = (x_n, y_n)$ ，则该通路是闭合通路。根据不同的邻接类型，可以得到不同的通路，如 4- 通路， 8- 通路和 m - 通路。

若像素 p 和 q 之间存在着一条通路，则称 p 和 q 是连通的，即 **像素间的连通**。反之，若两个像素间是连通的，那么至少存在一条通路，也可能存在多条通路。根据通路的类型，像素间的连通可分为 4- 连通， 8- 连通和 m - 连通。



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

➤ 像素间的距离

像素在空间上的接近程度用距离来度量。对于任意的像素 p , q 和 r , 坐标分别为 (x, y) 、 (s, t) 、 (u, v) , 若满足

- a) $D(p, q) \geq 0$, (当且仅当时等号成立);
- b) $D(p, q) = D(q, p)$;
- c) $D(p, r) \leq D(p, q) + D(q, r)$.

则称函数 D 为距离或度量函数。

在数字图像中, 对上述距离函数 D 有几种不同定义, 常见的有欧式距离、城区距离、棋盘距离等,



第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

① 欧氏距离

欧式距离是范数为2的距离，坐标为 (x, y) 的像素 p 和坐标为 (s, t) 的像素 q 之间的欧式距离定义如下：

$$D_e(p, q) = \sqrt{(x - s)^2 + (y - t)^2}$$

在这种情况下，距 p 的 D_e 距离小于或等于某一值 R 的像素都包含在以 p 为圆心且半径为 R 的圆内。



第1章 数字图像处理的基本知识

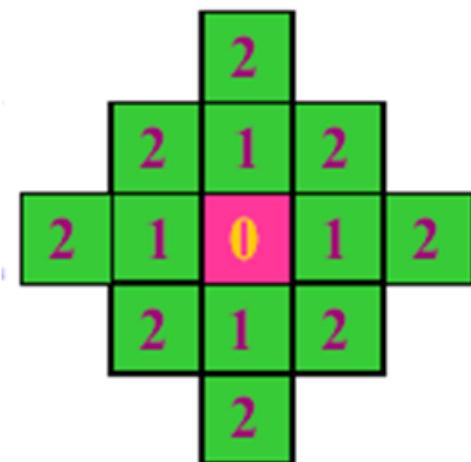
1.1 图像的概念、数字图像及表示

② 城区距离

城区距离是范数为1的距离，坐标为 (x, y) 的像素 p 和坐标为 (s, t) 的像素 q 之间的城区距离定义如下：

$$D_4(p, q) = |x - s| + |y - t|$$

在这种情况下，距 p 的 D_4 距离小于或等于某一值 R 的像素都包含在以 p 为中心的菱形区域。例如当 $R=2$ 时，菱形距离区域为：





第1章 数字图像处理的基本知识

1.1 图像的概念、数字图像及表示

③ 棋盘距离

棋盘距离是范数为 ∞ 的距离，坐标为 (x, y) 的像素 p 和坐标为 (s, t) 的像素 q 之间的棋盘距离定义如下：

$$D_8(p, q) = \max(|x - s|, |y - t|)$$

在这种情况下，距 p 的 D_8 距离小于或等于某一值 R 的像素都包含在以 p 为中心的正方形区域。例如当 $R=2$ 时，正方形距离区域为：

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2





第1章 数字图像处理的基本知识

1.2 数字图像的存储与读写

在数字图像处理系统中，图像文件数据量大，它们需要被存档以便以后的读写、处理及显示，而且经常需要在不同的用户和系统间交换。这就要求有一些用于数字图像文件存储和传送的标准格式。

图像数据文件的格式种类很多，不同的系统平台和软件使用不同的图像文件格式。

每一种图像文件均有一个文件头，在文件头之后是图像数据。文件头的内容一般包括文件类型、版本号、文件大小等内容。各种图像文件的制作还涉及到图像文件的压缩方式和存储效率等。常用的图像文件存储格式主要有： **BMP** 格式、 **JPEG** 格式、 **GIF** 格式、 **TIFF** 格式等。



第1章 数字图像处理的基本知识

1.2 数字图像的存储与读写

BMP(bitmap) 格式是微软公司为其 Windows 操作系统设置的标准图像格式，文件后缀名为“.bmp”，是个人计算机上最常见、最简单的文件格式之一。BMP 图像文件也称位图文件，其文件结构如下：

- a) BITMAPFILEHEADER : BMP 文件头或表头；
- b) BITMAPINFOHEADER : BMP 文件信息头；
- c) RGBQUAD : BMP 文件调色板或位图信息；
- d) BITMAP DATA : BMP 文件数据或位图阵列；

位图文件头长度为固定的 54 个字节，给出了文件的类型、大小和位图阵列的起始位置等信息。位图文件信息头基本上包含图像的所有信息，包括宽度、高度、每像素的位数、压缩方法、目标设备的水平和垂直分辨率等信息。位图文件调色板是 BMP 所包含的颜色表，接在 BITMAPINFOHEADER 结构之后含有位图中每种颜色的 RGB 信息。BMP 文件数据是以连续行的形式存储的。





文件部分	属性	说明
BTMAPFILEHEADER (位图文件头)	bfType	文件类型，必须是 0x424D,即字符串“BM”
	bfSize	指定文件大小，包括这 14 个字节
	bfReserved1	保留字，不用考虑
	bfReserved2	保留字，不用考虑
	bfOffBits	从文件头到实际位图数据的偏移字节数
BITMAPINFOHEADER (位图信息头)	biSize	该结构的长度，为 40
	biWidth	图像的宽度，单位是像素
	biHeight	图像的高度，单位是像素
	biPlanes	位平面数，必须是 1，不用考虑
	biBitCount	指定颜色位数，1 为二色，4 为 16 色，8 为 256 色，16、24、32 为真彩色
	biCompression	指定是否压缩，有效值为 BI_RGB、BI_RLE8、BI_RLE4、BI_BITFIELDS
	biSizeImage	实际的位图数据占用的字节数
	biXPelsPerMeter	目标设备水平分辨率，单位是每米的像素数
	biYPelsPerMeter	目标设备垂直分辨率，单位是每米的像素数
	biClrUsed	实际使用的颜色数，若该值为 0，则使用颜色数为 2 的 biBitCount 次方种
Palette (调色板)	rgbBlue	该颜色的蓝色分量
	rgbGreen	该颜色的绿色分量
	rgbRed	该颜色的红色分量
	rgbReserved	保留值
ImageData (位图数据)	像素按行优先顺序排列，每一行的字节数必须是 4 的整倍数	





第1章 数字图像处理的基本知识

1.2 数字图像的存储与读写

JPEG (Joint Photographic Experts Group) 的文件后缀名为“.jpg”或“.jpeg”。它由一个软件开发联合会组织制定，是一种有损压缩格式，能够将图像压缩在很小的存储空间。JPEG 格式支持 24 位颜色，并保留照片和其他连续色调图像中存在的亮度和色相的显著和细微的变化。

GIF (Graphics Interchange Format) 的文件后缀名为“.gif”。GIF 格式的特点是其在一个 GIF 文件中可以存多幅彩色图像，如果把存于一个文件中的多幅图像数据逐幅读出并显示到屏幕上，就可构成一种最简单的动画。

TIFF (Tag Image File Format) 的文件后缀名为“.tif”或“.tiff”。TIFF 文件一般可分为文件头、参数指针表、参数数据表的偏移。参数指针表由一系列参数块构成，它们描写图像的压缩种类、长度、彩色数、扫描分辨率等许多参数。参数数据表中存放的是实际参数数据。最后一部分是图像数据。



数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第1章 数字图象处理的基本知识

第3讲 数字图像文件的读写与显示





BMP 图像文件的读取、显示与存储 — VB





我的文档



我的电脑



网上邻居



回收站



Internet
Explorer



桌面图标



Microsoft
Visual Basic
6.0 中文版

开始



>>



文件(F) 编辑(E) 视图(V) 工程(P) 格式(O) 调试(U) 运行(R) 查询(I) 图表(L) 工具(T) 外接程序(A) 窗口(W) 帮助(H)



General



新建工程



确定

取消

帮助(H)



属性

按字母序 | 按分类序 |

窗体布局

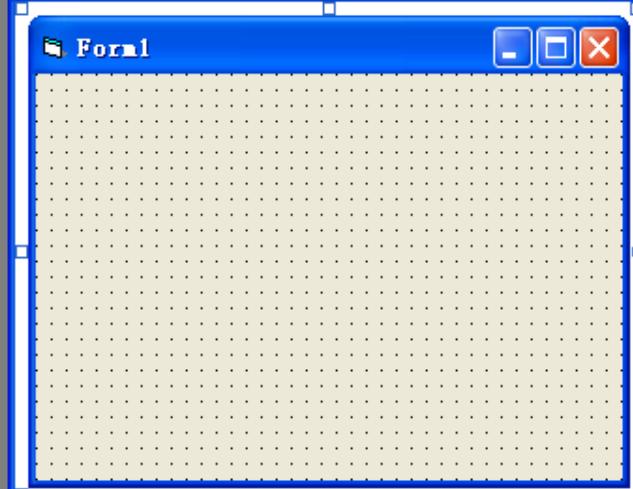


工程1 - Microsoft Visual Basic [设计]

文件(F) 编辑(E) 视图(V) 工程(P) 格式(O) 调试(D) 运行(R) 查询(U) 图表(I) 工具(T) 外接程序(A) 窗口(W) 帮助(H)



工程1 - Form1 (Form)



0, 0 4815 x 3645

工程 - 工程1



属性 - Form1

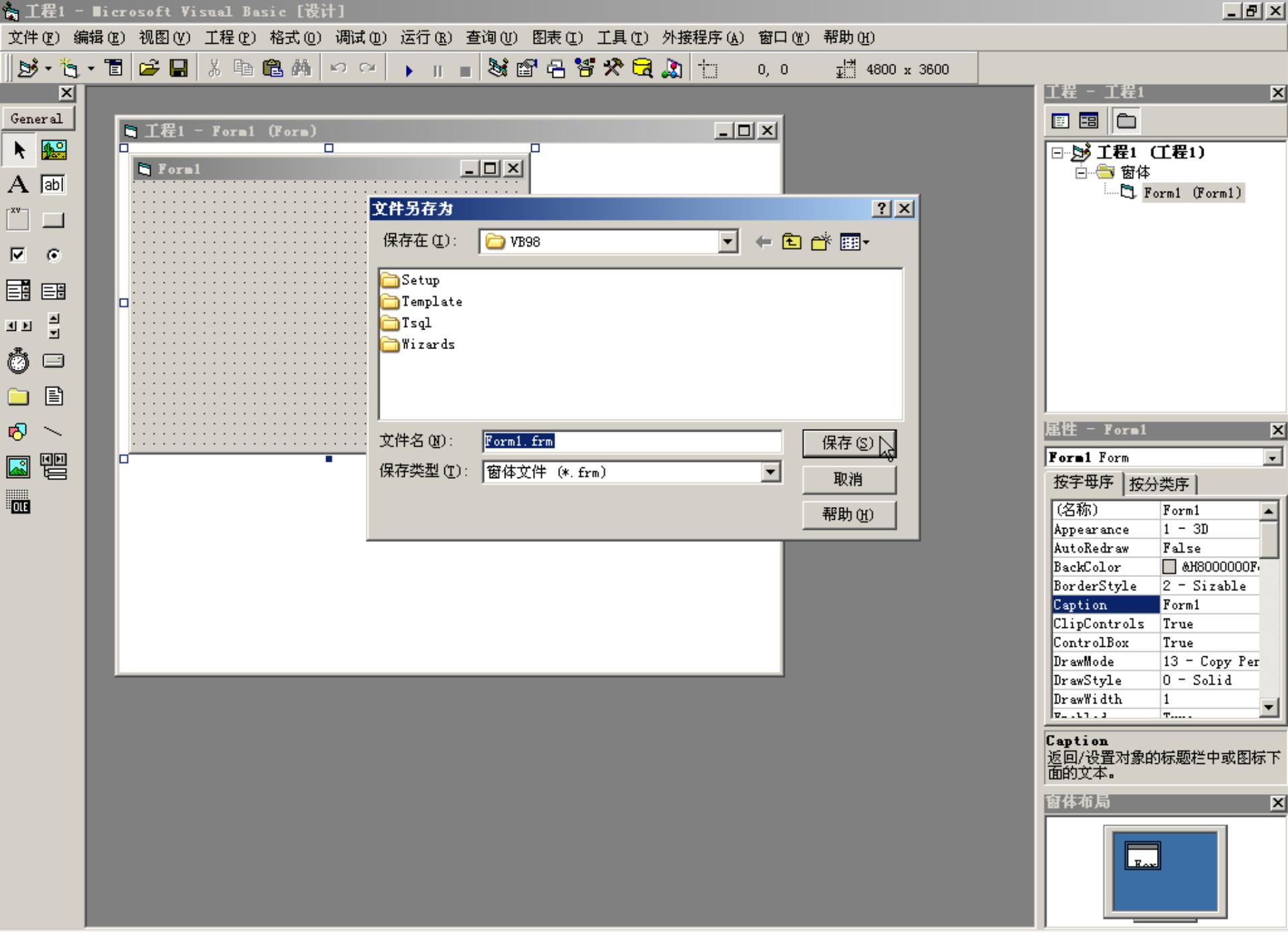
Form1 Form	
按字母序 按分类序	
(名称)	Form1
Appearance	1 - 3D
AutoRedraw	False
BackColor	0 &H8000000F
BorderStyle	2 - Sizable
Caption	Form1
ClipControls	True
ControlBox	True
DrawMode	13 - Copy Per
DrawStyle	0 - Solid
DrawWidth	1

Caption

返回/设置对象的标题栏中或图标下面的文本。

窗体布局





工程1 - Microsoft Visual Basic [设计]

文件(F) 编辑(E) 视图(V) 工程(P) 格式(O) 调试(D) 运行(R) 查询(U) 图表(I) 工具(T) 外接程序(A) 窗口(W) 帮助(H)



General



工程1 - Form1 (Form)

Form1

0, 0 4815 x 3645



属性 - Form1

Form1 Form	
	按字母序 按分类序
(名称)	Form1
Appearance	1 - 3D
AutoRedraw	False
BackColor	□ &H8000000F
BorderStyle	2 - Sizable
Caption	Form1
ClipControls	True
ControlBox	True
DrawMode	13 - Copy Per
DrawStyle	0 - Solid
DrawWidth	1

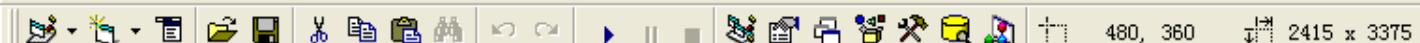
Caption
返回/设置对象的标题栏中或图标下面的文本。



工程1 - Microsoft Visual Basic [设计]



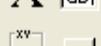
文件(F) 编辑(E) 视图(V) 工程(P) 格式(O) 调试(U) 运行(R) 查询(U) 图表(I) 工具(T) 外接程序(A) 窗口(W) 帮助(H)



General



A ab



✓ ◎



S S



Clock



File



Form



OLE

工程1 - Form1 (Form)



Form1



工程 - 工程1



工程1 (实验一.vbp)

窗体

Form1 (实验一.frm)

属性 - Picture1

Picture1 PictureBox

按字母序 | 按分类序

Left	480
LinkItem	
LinkMode	0 - None
LinkTimeout	50
LinkTopic	
MouseIcon	(None)
MousePointer	0 - Default
Negotiate	False
OLEDragMode	0 - Manual
OLEDropMode	0 - None
Picture	(None)

Picture

返回/设置控件中显示的图形。

窗体布局



工程1 - Microsoft Visual Basic [设计]



文件(F) 编辑(E) 视图(V) 工程(P) 格式(O) 调试(U) 运行(R) 查询(U) 图表(I) 工具(T) 外接程序(A) 窗口(W) 帮助(H)

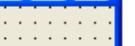
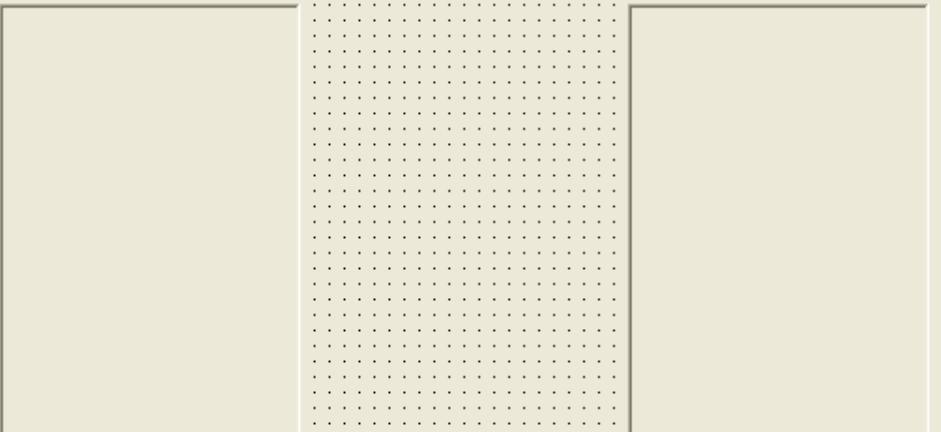


General



工程1 - Form1 (Form)

Form1



工程 - 工程1



工程1 (实验一.vbp)



窗体



Form1 (实验一.frm)

属性 - Form1

Form1 Form

按字母序 | 按分类序

MaxButton	True
MDIChild	False
MinButton	True
MouseIcon	(None)
MousePointer	0 - Default
Moveable	True
NegotiateMenus	True
OLEDropMode	0 - None
Palette	(None)
PaletteMode	0 - Halftone
Picture	(None)

Picture

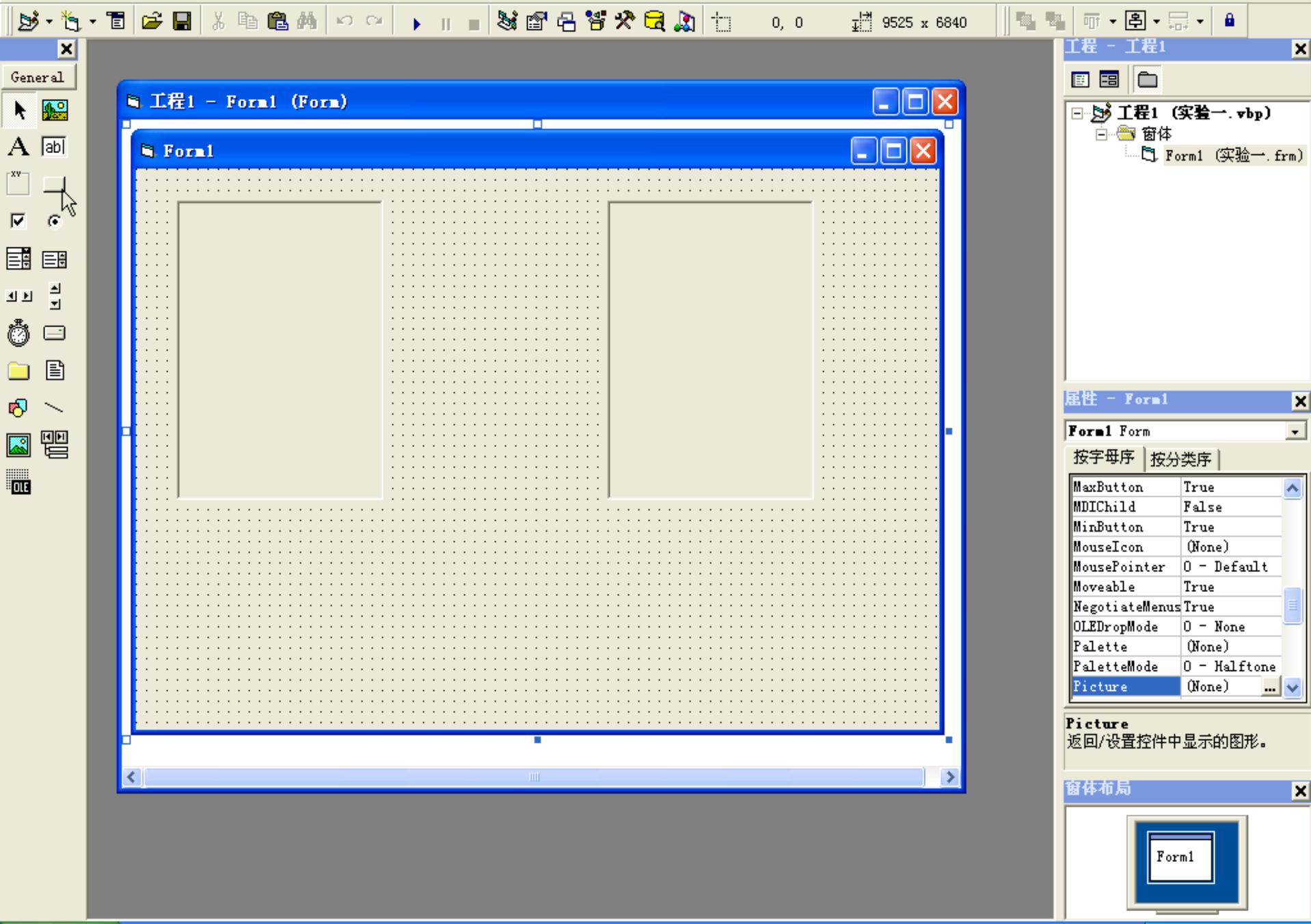
返回/设置控件中显示的图形。

窗体布局



工程1 - Microsoft Visual Basic [设计]

文件(F) 编辑(E) 视图(V) 工程(P) 格式(O) 调试(U) 运行(R) 查询(U) 图表(I) 工具(T) 外接程序(A) 窗口(W) 帮助(H)



工程1 - Microsoft Visual Basic [设计]



文件(F) 编辑(E) 视图(V) 工程(P) 格式(O) 调试(D) 运行(R) 查询(U) 图表(I) 工具(T) 外接程序(A) 窗口(W) 帮助(H)



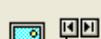
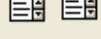
General



A ab



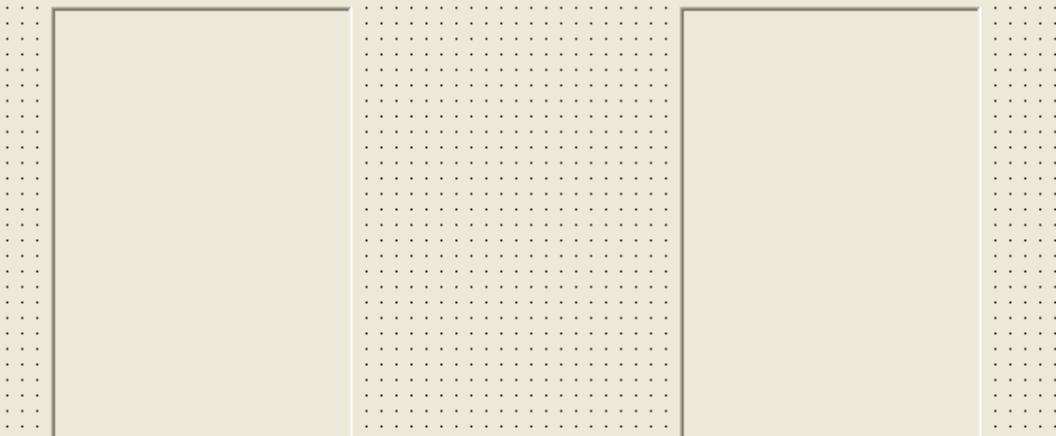
✓ ◎



工程1 - Form1 (Form)



Form1



工程 - 工程1

工程1 (实验一.vbp)

窗体
Form1 (实验一.frm)

属性 - Command1

Command1 CommandButton

按字母序 | 按分类序

(名称)	Command1
Appearance	1 - 3D
BackColor	0x0000000F
Cancel	False
Caption	Command1
CausesValidation	True
Default	False
DisabledPicture	(None)
DownPicture	(None)
DragIcon	(None)
DragMode	0 - Manual

(名称)

返回代码中使用的标识对象的名称。

窗体布局



工程1 - Microsoft Visual Basic [设计]



文件(F) 编辑(E) 视图(V) 工程(P) 格式(O) 调试(D) 运行(R) 查询(U) 图表(I) 工具(T) 外接程序(A) 窗口(W) 帮助(H)



General



工程1 - Form1 (Form)

Form1

Command1 Command2 Command5 Command3 Command4



属性 - Form1

Form1 Form

	按字母序	按分类序
(名称)	Form1	
Appearance	1 - 3D	
AutoRedraw	False	
BackColor	0x0000000F	
BorderStyle	2 - Sizable	
Caption	Form1	
ClipControls	True	
ControlBox	True	
DrawMode	13 - Copy Per	
DrawStyle	0 - Solid	
DrawWidth	1	

Caption
返回/设置对象的标题栏中或图标下面的文本。



工程1 - Microsoft Visual Basic [设计]

文件(F) 编辑(E) 视图(V) 工程(P) 格式(O) 调试(U) 运行(R) 查询(U) 图表(I) 工具(T) 外接程序(A) 窗口(W) 帮助(H)

The screenshot shows the Microsoft Visual Basic 6.0 IDE interface. The main window displays the code for Form1 (Form). The code window title is "工程1 - Form1 (Code)" and contains the following VBA code:

```
Private Sub Form_Load()
End Sub
```

The properties window on the right, titled "属性 - Form1", lists the properties for Form1. The "Caption" property is currently selected and set to "Form1". Other properties include:

(名称)	Form1
Appearance	1 - 3D
AutoRedraw	False
BackColor	0x000000FF
BorderStyle	2 - Sizable
Caption	Form1
ClipControls	True
ControlBox	True
DrawMode	13 - Copy Per
DrawStyle	0 - Solid
DrawWidth	1

The status bar at the bottom indicates "行 2, 列 5". The toolbar on the left contains various icons for file operations, such as Open, Save, Print, and Find.

工程1 - Microsoft Visual Basic [设计]

文件(F) 编辑(E) 视图(V) 工程(P) 格式(O) 调试(U) 运行(R) 查询(U) 图表(I) 工具(T) 外接程序(A) 窗口(W) 帮助(H)

The screenshot shows the Microsoft Visual Basic 6.0 IDE interface. The main window is titled "工程1 - Form1 (Form)" and contains a smaller window titled "工程1 - Form1 (Code)". The "Code" window displays the following VBA code:

```
Private Sub Form_Load()
End Sub
```

The "Properties" window on the right lists the properties for "Form1 Form". The "Caption" property is currently selected and set to "Form1". The "Caption" section of the status bar at the bottom provides a brief description of the property.

(名称)	Form1
Appearance	1 - 3D
AutoRedraw	False
BackColor	0 &H8000000F
BorderStyle	2 - Sizable
Caption	Form1
ClipControls	True
ControlBox	True
DrawMode	13 - Copy Per
DrawStyle	0 - Solid
DrawWidth	1

Caption
返回/设置对象的标题栏中或图标下面的文本。

属性 - Form1

按字母序 | 按分类序

Form1 Form

Caption

返回/设置对象的标题栏中或图标下面的文本。

窗体布局

Form1

General

Form

Load

```
Private Sub Form_Load()
Picture1.AutoScale = True      '设置Picture1的大小可以随读入图片进行变化，使其与图片的大小相符
Picture2.AutoRedraw = True    '设置picture2可进行重画
Command1.Caption = "打开图像"
Command2.Caption = "保存图像"
Command3.Caption = "图像灰度化"
Command4.Caption = "保存灰度图像"
Command5.Caption = "退出"
End Sub
```

属性

按字母序 | 按分类序

窗体布局

工程 - 工程1

工程1 (实验一.vbp)

窗体

Form1 (实验一.frm)

窗体

文件(F) 编辑(E) 视图(V) 工程(P) 格式(O) 调试(R) 运行(R) 查询(U) 图表(I) 工具(T) 外接程序(A) 窗口(W) 帮助(H)



Form

Load

General

```
Private Sub Form_Load()
Picture1.AutoScale = True      '设置Picture1的大小可以随读入图片进行变化，使其与图片的大小相符
Picture2.AutoRedraw = True    '设置picture2可进行重画
Command1.Caption = "打开图像"
Command2.Caption = "保存图像"
Command3.Caption = "图像灰度化"
Command4.Caption = "保存灰度图像"
Command5.Caption = "退出"
End Sub
```

工程 - 工程1

工程1 (实验一.vbp)

窗体

Form1 (实验一.frm)

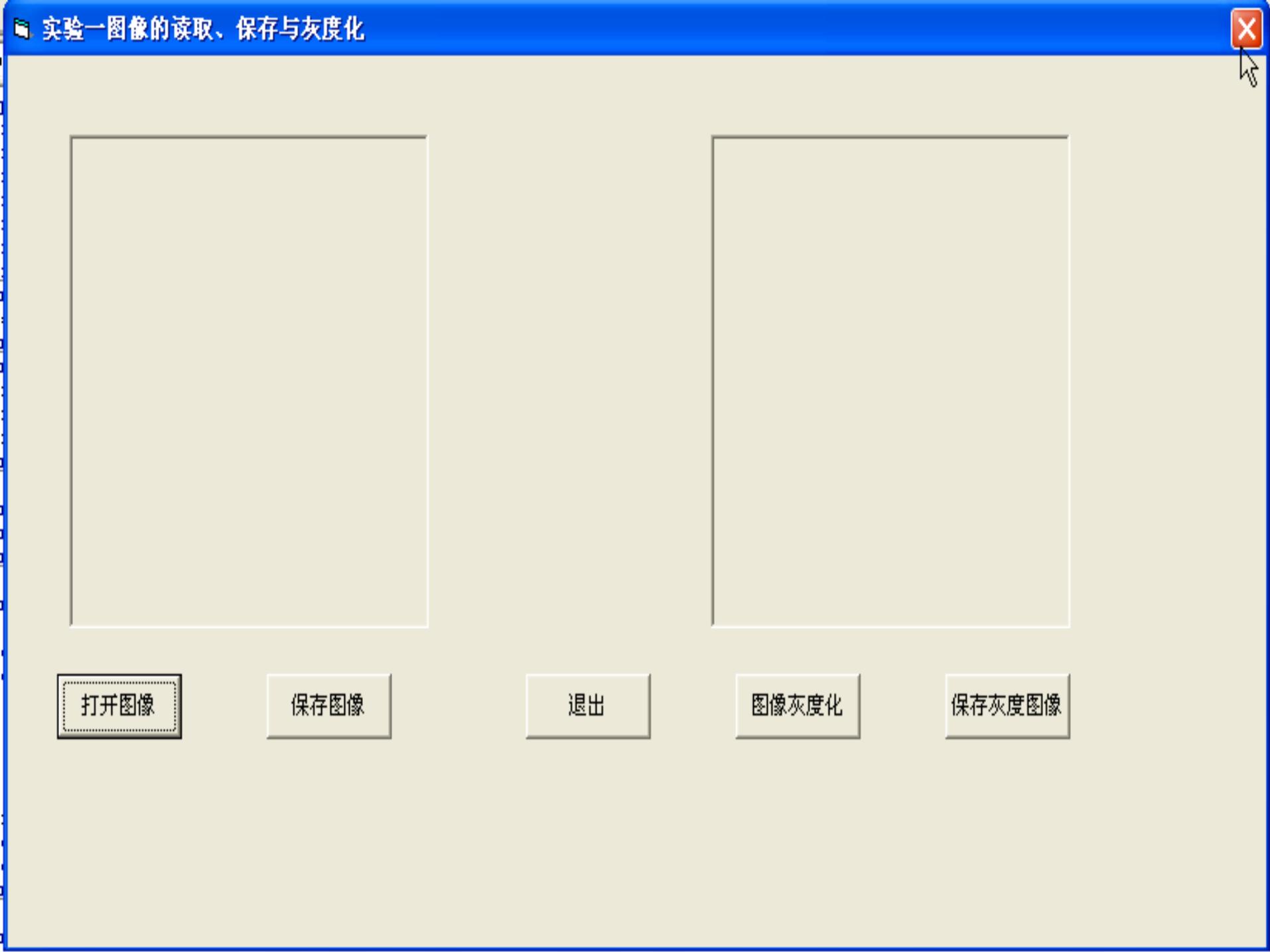
属性

按字母序 | 按分类序

窗体布局

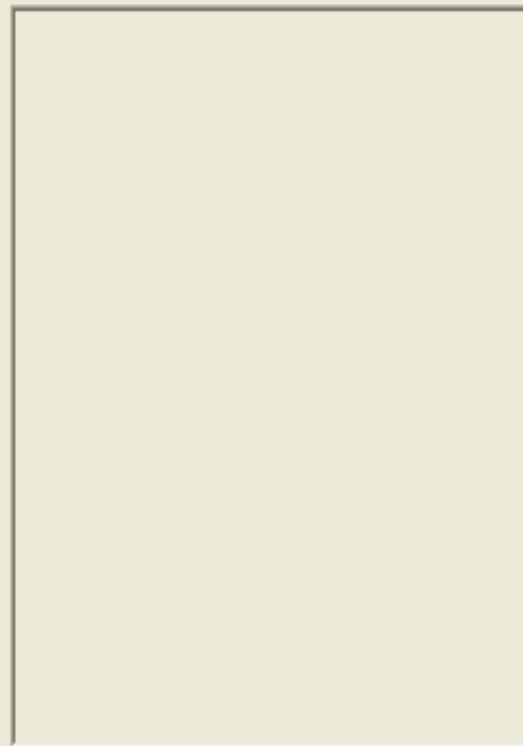
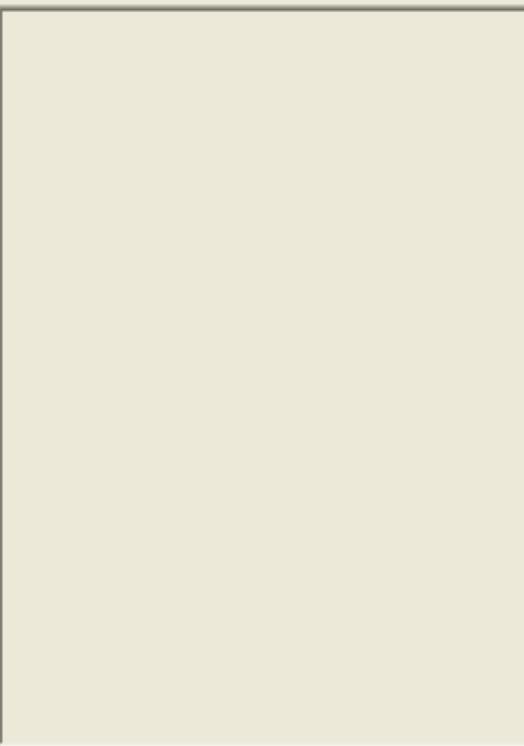
OLE

This screenshot shows the Microsoft Visual Basic 6.0 IDE interface during design mode. The main code editor window displays VBA code for the Form1_Load event. The code initializes two picture boxes (Picture1 and Picture2) and five command buttons (Command1 to Command5) with specific captions. A note in the code explains the purpose of setting Picture1.AutoScale to True. The project tree on the right shows the '工程1' folder containing '窗体' and 'Form1 (实验一.frm)'. Below the project tree is the '属性' (Properties) panel, which is currently empty. At the bottom right is the '窗体布局' (Form Layout) panel, showing a preview of the form with a single blue square placeholder. The left side features the standard Windows-style toolbar with icons for file operations, selection tools, and other common functions.



实验一图像的读取、保存与灰度化

X



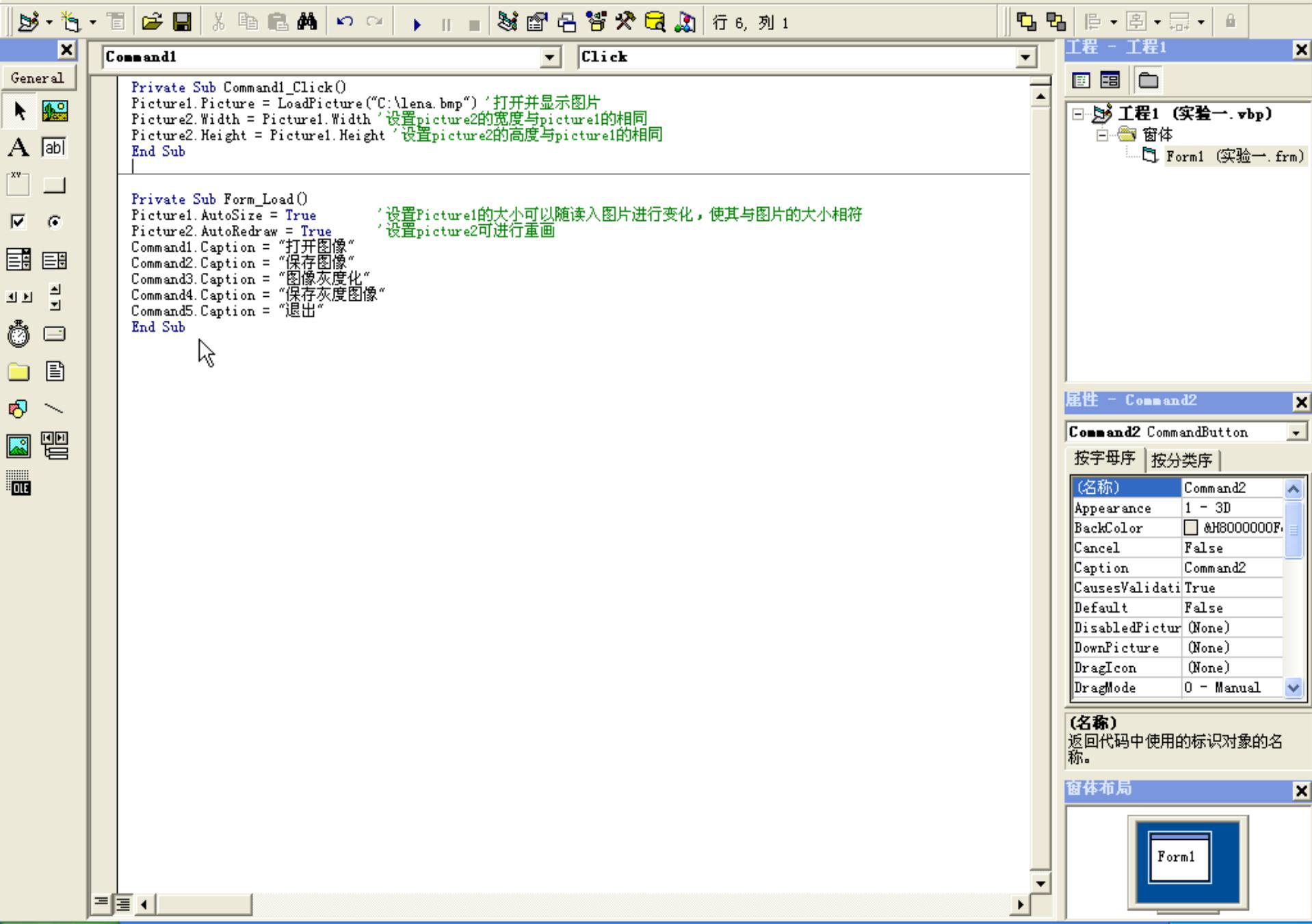
打开图像

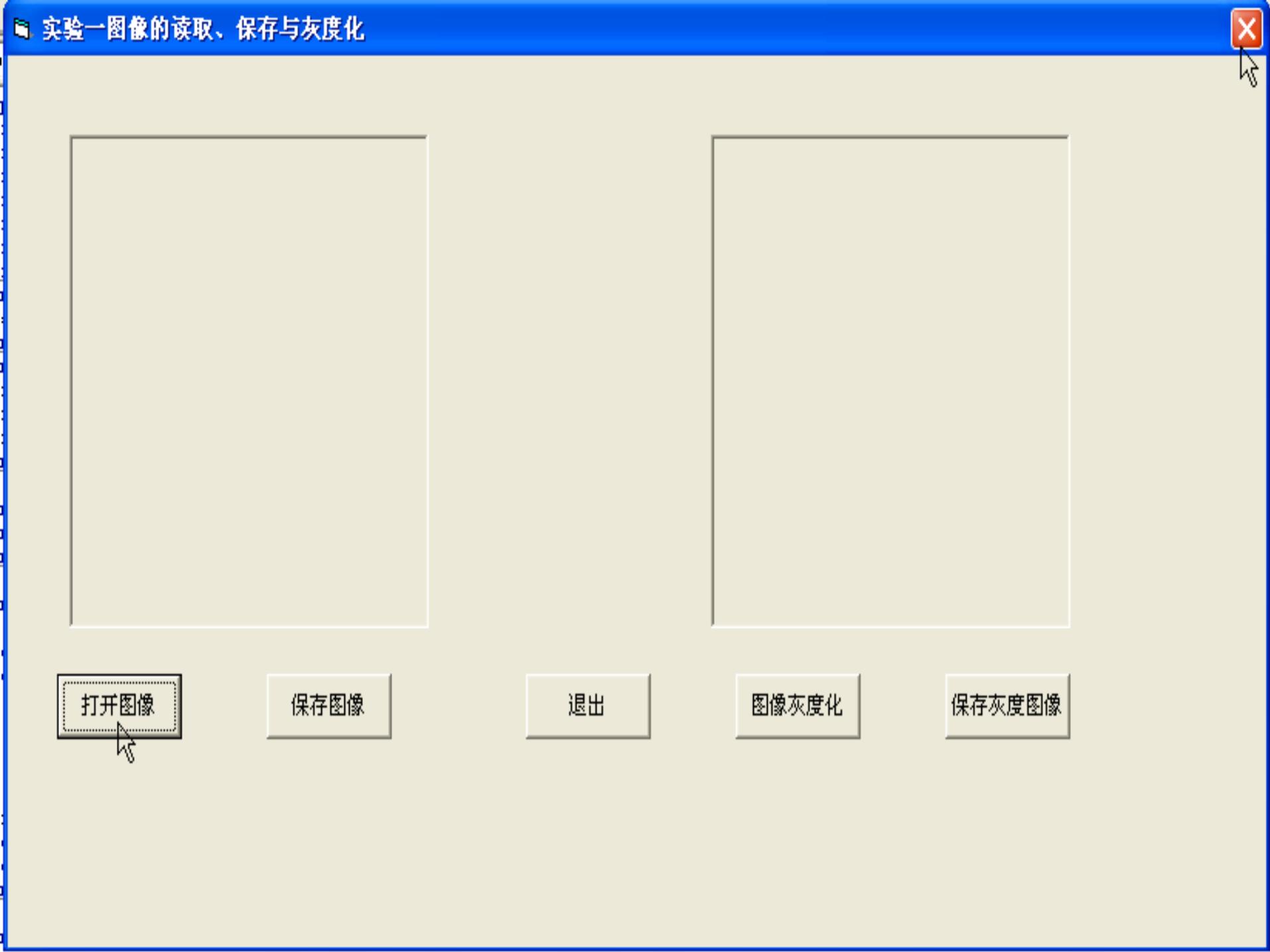
保存图像

退出

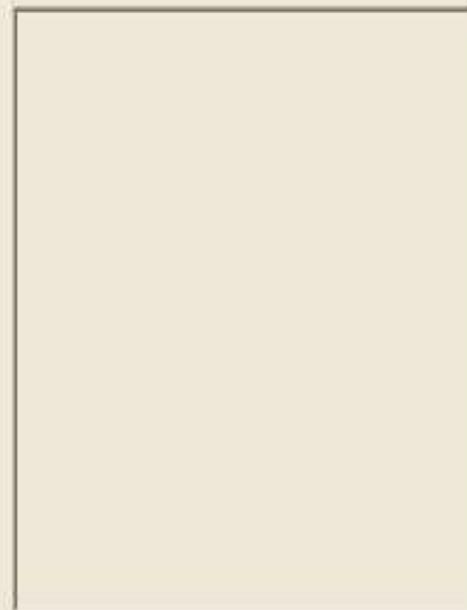
图像灰度化

保存灰度图像





实验一图像的读取、保存与灰度化



打开图像

保存图像

退出

图像灰度化

保存灰度图像

行 8, 列 43

Command2 Click

```

Private Sub Command1_Click()
Picture1.Picture = LoadPicture("C:\lena.bmp") ' 打开并显示图片
Picture2.Width = Picture1.Width ' 设置picture2的宽度与picture1的相同
Picture2.Height = Picture1.Height ' 设置picture2的高度与picture1的相同
End Sub

Private Sub Command2_Click()
SavePicture Picture1.Picture, "C:\new.bmp"
End Sub

Private Sub Form_Load()
Picture1.AutoSize = True      ' 设置Picture1的大小可以随读入图片进行变化，使其与图片的大小相符
Picture2.AutoRedraw = True   ' 设置picture2可进行重画
Command1.Caption = "打开图像"
Command2.Caption = "保存图像"
Command3.Caption = "图像灰度化"
Command4.Caption = "保存灰度图像"
Command5.Caption = "退出"
End Sub

```

工程 - 工程1

- 工程1 (实验一.vbp)
 - 窗体
 - Form1 (实验一.frm)

属性 - Command2

按字母序 按分类序	
(名称)	Command2
Appearance	1 - 3D
BackColor	<input type="color"/>
Cancel	False
Caption	Command2
CausesValidation	True
Default	False
DisabledPicture	(None)
DownPicture	(None)
DragIcon	(None)
DragMode	0 - Manual

(名称)
返回代码中使用的标识对象的名称。

窗体布局

General

命令按钮

文本框

复选框

单选按钮

列表框

组合框

进度条

时钟

文件夹

图标

OLE

Image

Form

Command3

Click

```

Private Sub Command1_Click()
Picture1.Picture = LoadPicture("C:\lena.bmp") '打开并显示图片
Picture2.Width = Picture1.Width '设置picture2的宽度与picture1的相同
Picture2.Height = Picture1.Height '设置picture2的高度与picture1的相同
End Sub

Private Sub Command2_Click()
SavePicture Picture1.Picture, "C:\new.bmp"
End Sub

Private Sub Command3_Click()
For i = 0 To Picture1.Width - 1
For j = 0 To Picture1.Height - 1
c = Picture1.Point(i, j)
R = c And &HFF&
G = (c And &HFF00) / 256
B = (c And &HFF0000) / 65536
y = (R + G + B) / 3
Picture2.PSet (i, j), RGB(y, y, y)
Next j
Next i
End Sub

Private Sub Form_Load()
Picture1.AutoSize = True      '设置Picture1的大小可以随读入图片进行变化，使其与图片的大小相符
Picture2.AutoRedraw = True   '设置picture2可进行重画
Command1.Caption = "打开图像"
Command2.Caption = "保存图像"
Command3.Caption = "图像灰度化"
Command4.Caption = "保存灰度图像"
Command5.Caption = "退出"
End Sub

```

工程 - 工程1

- 工程1 (实验一.vbp)
 - 窗体
 - Form1 (实验一.frm)

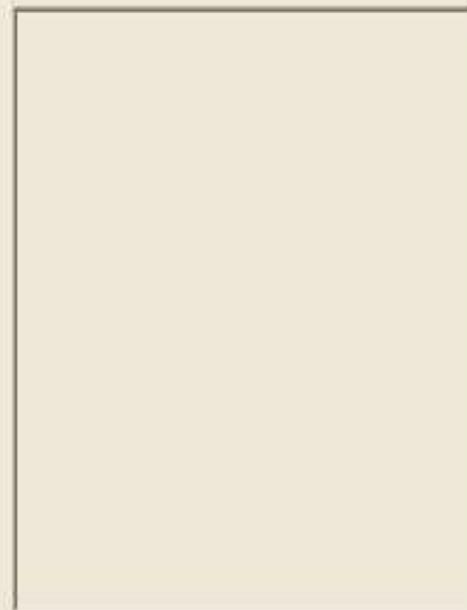
属性 - Command3

Command3 CommandButton	
	按字母序 按分类序
(名称)	Command3
Appearance	1 - 3D
BackColor	□ &H8000000F
Cancel	False
Caption	Command3
CausesValidation	True
Default	False
DisabledPicture	(None)
DownPicture	(None)
DragIcon	(None)
DragMode	0 - Manual

(名称)
返回代码中使用的标识对象的名称。

窗体布局

实验一图像的读取、保存与灰度化



打开图像

保存图像

退出

图像灰度化

保存灰度图像

实验一图像的读取、保存与灰度化

X



打开图像

保存图像

退出

图像灰度化

保存灰度图
像

行 25, 列 44

Command4

Click

```

Private Sub Command1_Click()
Picture1.Picture = LoadPicture("C:\lena.bmp") ' 打开并显示图片
Picture2.Width = Picture1.Width ' 设置picture2的宽度与picture1的相同
Picture2.Height = Picture1.Height ' 设置picture2的高度与picture1的相同
End Sub

Private Sub Command2_Click()
SavePicture Picture1.Picture, "C:\new.bmp"
End Sub

Private Sub Command3_Click()
For i = 0 To Picture1.Width - 1
For j = 0 To Picture1.Height - 1
c = Picture1.Point(i, j)
R = c And &HFF&
G = (c And &HFF000) / 256
B = (c And &HFF0000) / 65536
y = (R + G + B) / 3
Picture2.PSet (i, j), RGB(y, y, y)
Next j
Next i
End Sub

Private Sub Command4_Click()
SavePicture Picture2.Image, "C:\huihua.bmp"
End Sub

Private Sub Form_Load()
Picture1.AutoSize = True      ' 设置Picture1的大小可以随读入图片进行变化，使其与图片的大小相符
Picture2.AutoRedraw = True    ' 设置picture2可进行重画
Command1.Caption = "打开图像"
Command2.Caption = "保存图像"
Command3.Caption = "图像灰度化"
Command4.Caption = "保存灰度图像"
Command5.Caption = "退出"
End Sub

```

工程 - 工程1

- 工程1 (实验一.vbp)
 - 窗体
 - Form1 (实验一.frm)

属性 - Command4

Command4 CommandButton	
按字母序 按分类序	
(名称)	Command4
Appearance	1 - 3D
BackColor	□ &H8000000F
Cancel	False
Caption	Command4
CausesValidation	True
Default	False
DisabledPicture	(None)
DownPicture	(None)
DragIcon	(None)
DragMode	0 - Manual

(名称)
返回代码中使用的标识对象的名称。

窗体布局

General

命令按钮

文本框

复选框

单选按钮

列表框

组合框

进度条

时钟

文件夹

图标

OLE

Command4

Click

```

Private Sub Command1_Click()
Picture1.Picture = LoadPicture("C:\lena.bmp") '打开并显示图片
Picture2.Width = Picture1.Width '设置picture2的宽度与picture1的相同
Picture2.Height = Picture1.Height '设置picture2的高度与picture1的相同
End Sub

Private Sub Command2_Click()
SavePicture Picture1.Picture, "C:\new.bmp"
End Sub

Private Sub Command3_Click()
For i = 0 To Picture1.Width - 1
For j = 0 To Picture1.Height - 1
c = Picture1.Point(i, j)
R = c And &HFF&
G = (c And &HFF00&) / 256
B = (c And &HFF0000) / 65536
y = (R + G + B) / 3
Picture2.PSet (i, j), RGB(y, y, y)
Next j
Next i
End Sub

Private Sub Command4_Click()
SavePicture Picture2.Image, "C:\huihua.bmp"
End Sub

Private Sub Command5_Click()
End
End Sub

Private Sub Form_Load()
Picture1.AutoSize = True      '设置Picture1的大小可以随读入图片进行变化，使其与图片的大小相符
Picture2.AutoRedraw = True    '设置picture2可进行重画
Command1.Caption = "打开图像"
Command2.Caption = "保存图像"
Command3.Caption = "图像灰度化"
Command4.Caption = "保存灰度图像"
Command5.Caption = "退出"
End Sub

```

工程 - 工程1

- 工程1 (实验一.vbp)
 - 窗体
 - Form1 (实验一.frm)

属性 - Command5

Command5 CommandButton	
	按字母序 按分类序
(名称)	Command5
Appearance	1 - 3D
BackColor	□ &H8000000F
Cancel	False
Caption	Command5
CausesValidation	True
Default	False
DisabledPicture	(None)
DownPicture	(None)
DragIcon	(None)
DragMode	0 - Manual

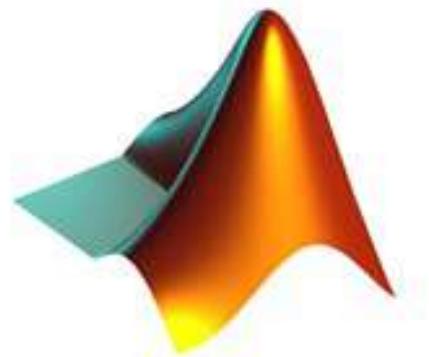
(名称)
返回代码中使用的标识对象的名称。

窗体布局





BMP 图像文件的读取、显示与存储 — MATLAB



MATLAB 7.0

MATLAB

File Edit Debug Desktop Window Help

Current Directory: D:\matlab\work

Shortcuts How to Add What's New

Workspace		
Name	Value	Class

Command Window

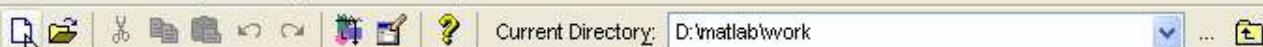
```
>> |
```

Start

MATLAB



File Edit Debug Desktop Window Help



Current Directory: D:\matlab\work



S New M-File How to Add What's New



Command Window

>>

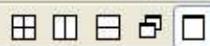


Editor - Untitled



File Edit Text Cell Tools Debug Desktop Window Help

~ | > x



1

Stack: Base

script

Ln 1 Col 1

OVR

```
1 i=imread('D:\matlab图像读取级灰度变换\lena.bmp'); %读取图像
2 figure, imshow(i); %显示图像
```

Editor - Untitled*

File Edit Text Cell Tools Debug Desktop Window Help

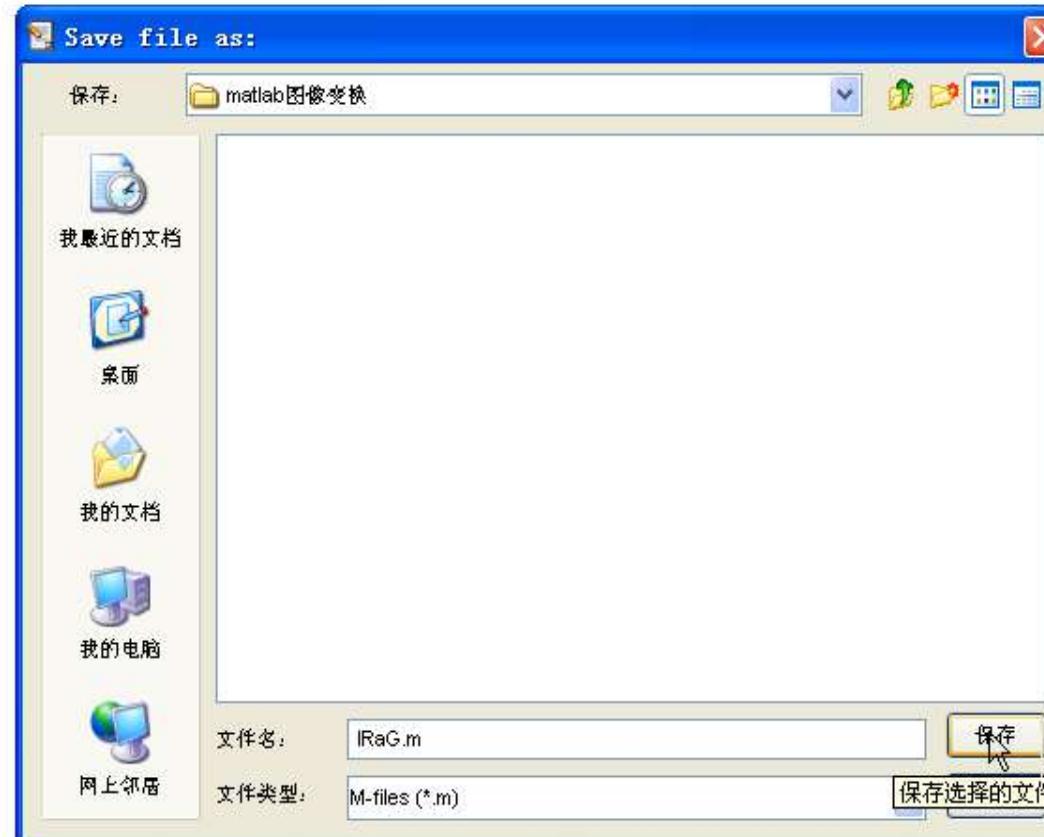
Save and run Stack: Base

```
1 i=imread('D:\matlab图像读取级联菜单\lena.bmp'); %读取图像
2 figure, imshow(i); %显示图像
```

script Ln 2 Col 32 OVR



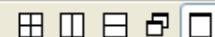
```
1 i=imread('D:\matlab图像读取级灰度变换\lena.bmp'); %读取图像  
2 figure, imshow(i); %显示图像
```



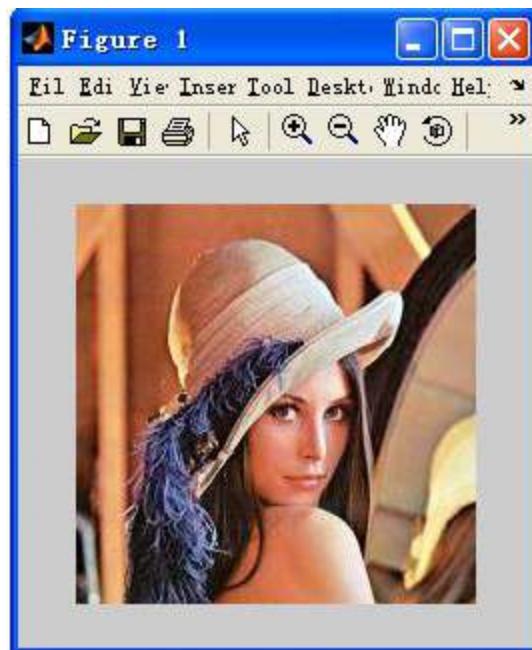
File Edit Text Cell Tools Debug Desktop Window Help



Stack: Base



```
1 - i=imread('D:\matlab图像读取级灰度变换\lena.bmp'); %读取图像  
2 - figure, imshow(i); %显示图像
```



File Edit Text Cell Tools Debug Desktop Window Help

Stack: Base

```
1 - i=imread('D:\matlab图像读取级灰度变换\lena.bmp'); %读取图像
2 - figure, imshow(i); %显示图像
3 - imwrite(i, '01.bmp'); %存储图像
```



Stack: Base

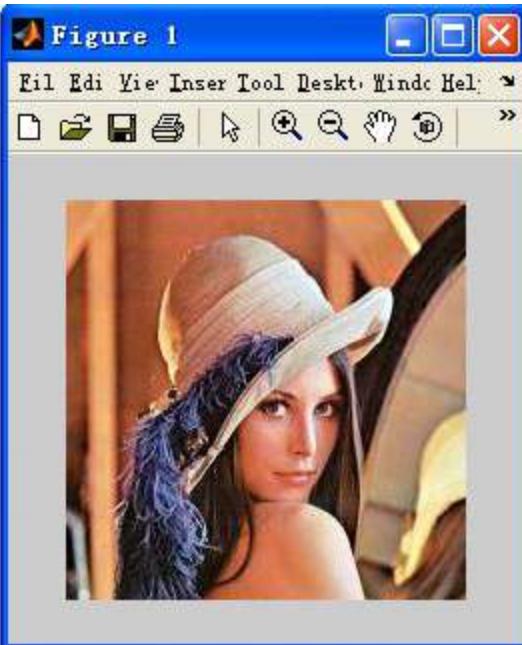
```
1 - i=imread('D:\matlab图像读取级灰度变换\lena.bmp'); %读取图像
2 - figure, imshow(i); %显示图像
3 - imwrite(i, '01.bmp'); %存储图像
4
5 - j=rgb2gray(i); %rgb图像转为灰度图像
6 - figure, imshow(j); %显示灰度图像
```

File Edit Text Cell Tools Debug Desktop Window Help

| | | | Stack: Base

```
1 - i=imread('D:\matlab图像读取级次\lena.bmp'); %读取图像
2 - figure, imshow(i); %显示图像
3 - imwrite(i, '01.bmp'); %存储图像
4
5 - j=rgb2gray(i); %rgb图像转为灰度图像
6 - figure, imshow(j); %显示灰度图像
```

```
1 - i=imread('D:\matlab图像读取级灰度变换\lena.bmp'); %读取图像  
2 - figure, imshow(i); %显示图像  
3 - imwrite(i, '01.bmp'); %存储图像  
4  
5 - j=rgb2gray(i); %rgb图像转为灰度图像  
6 - figure, imshow(j); %显示灰度图像
```





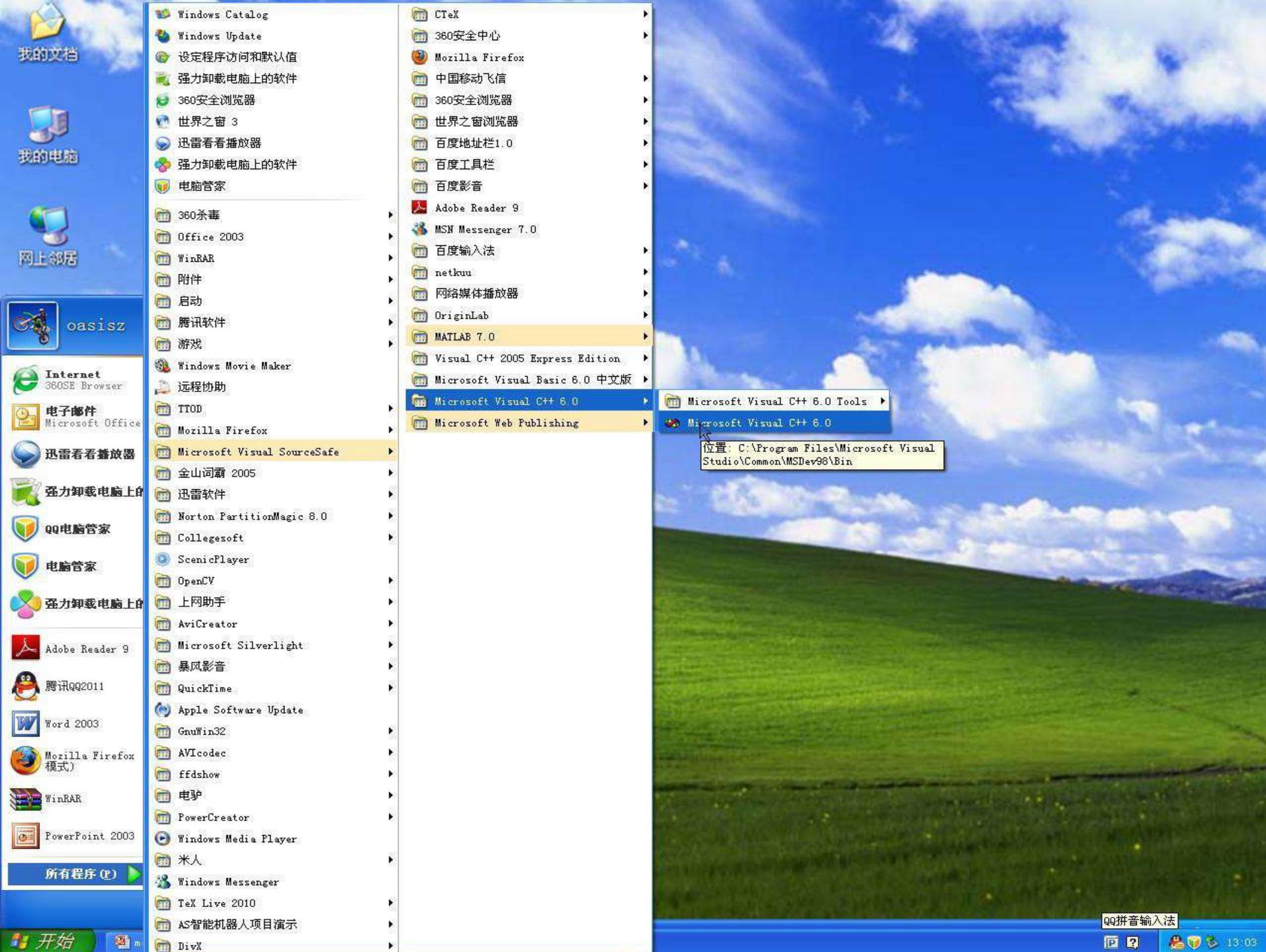
Stack: Base



```
1 - i=imread('D:\matlab图像读取级灰度变换\lena.bmp'); %读取图像
2 - figure, imshow(i); %显示图像
3 - imwrite(i, '01.bmp'); %存储图像
4
5 - j=rgb2gray(i); %rgb图像转为灰度图像
6 - figure, imshow(j); %显示灰度图像
7 - imwrite(j, '02.bmp'); %存储灰度图像
```



BMP 图像文件的读取、显示与存储 — VC





文件(F) 编辑(E) 查看(V) 插入(I) 工程(E) 组建(B) 工具(T) 窗口(W) 帮助(H)



就绪

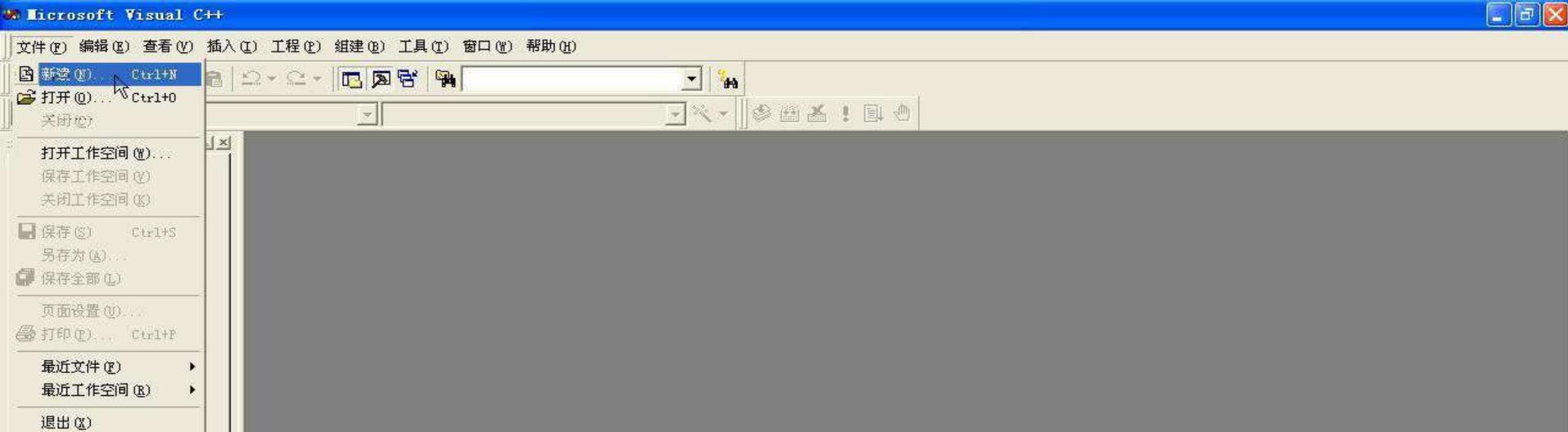
qq拼音输入法

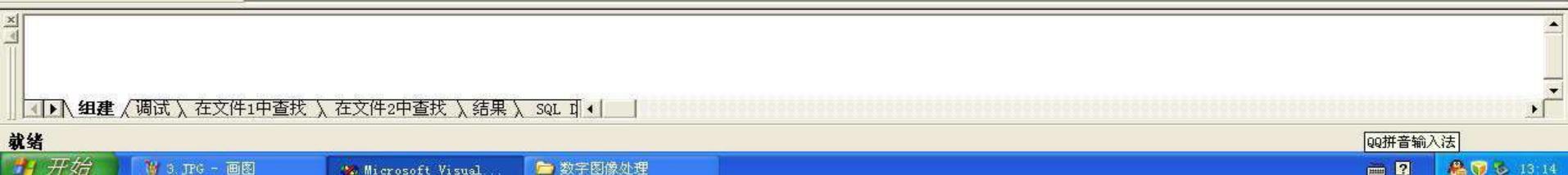
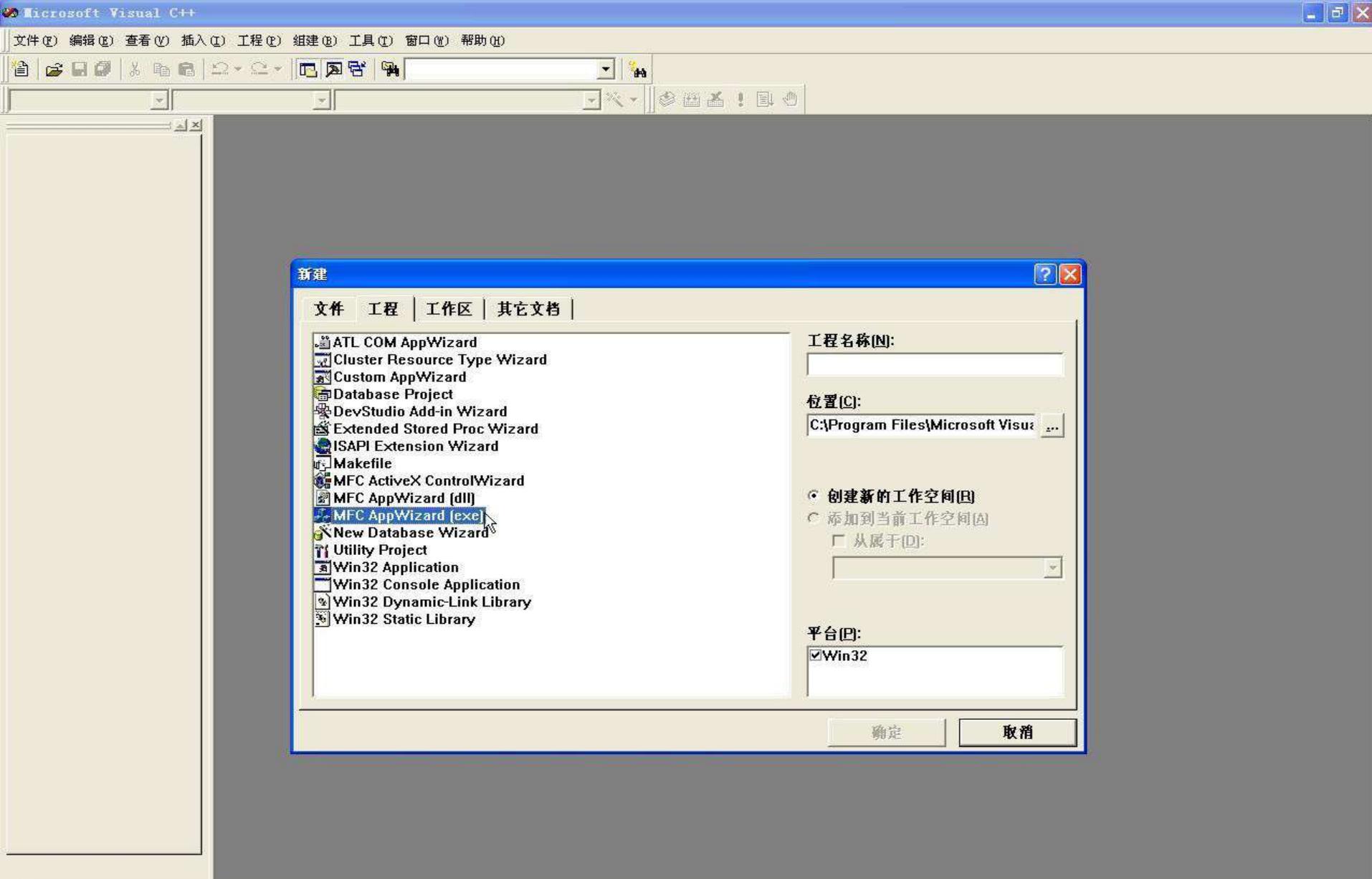


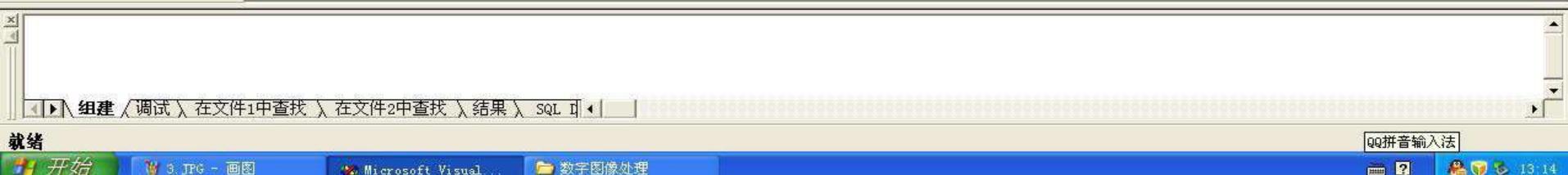
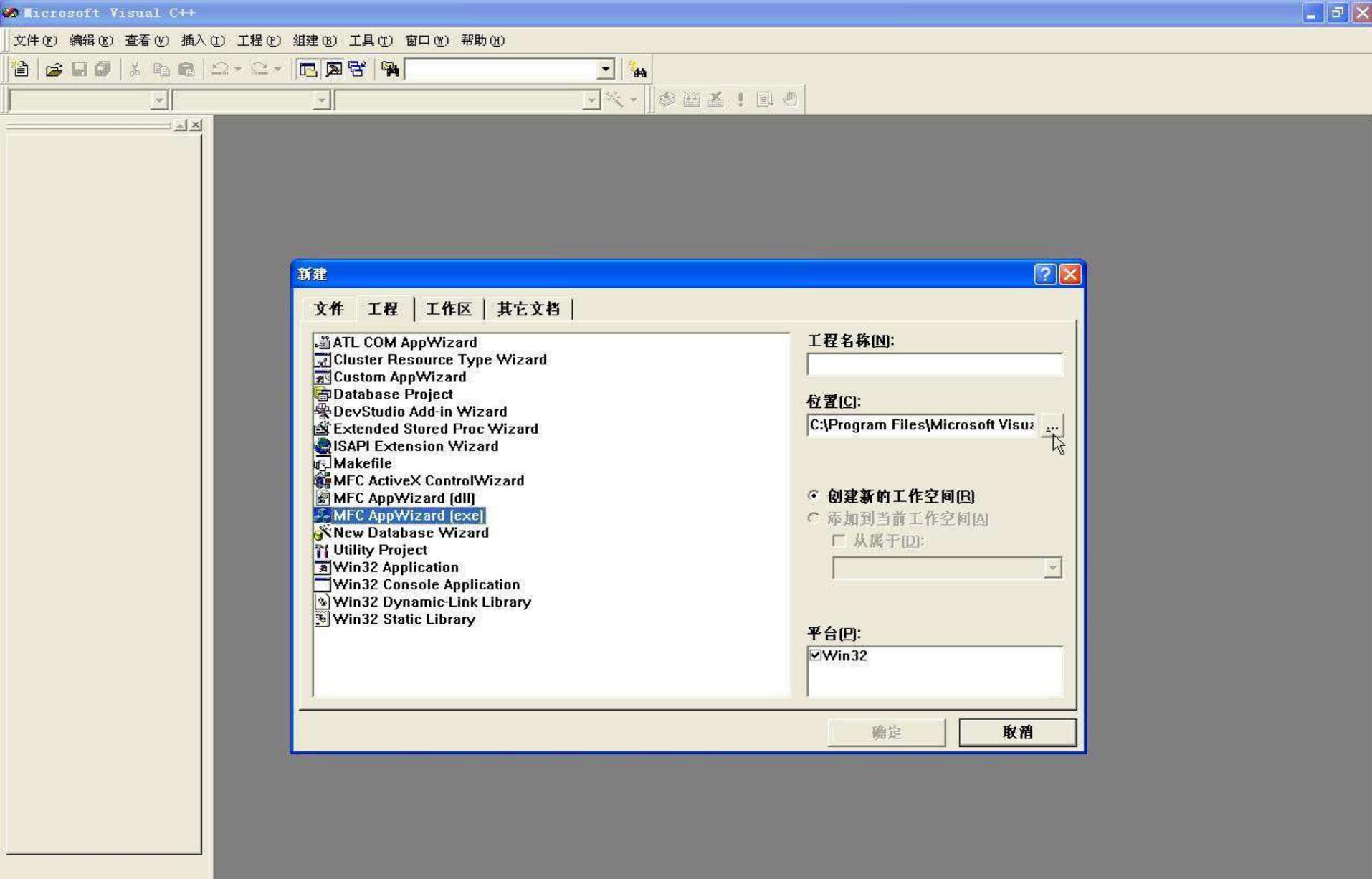
1.JPG - 画图

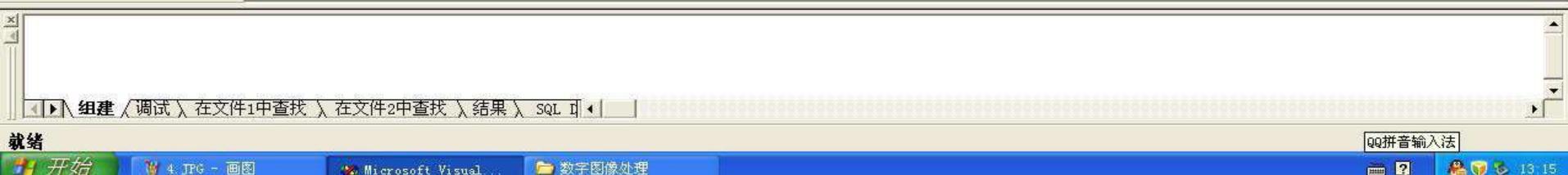
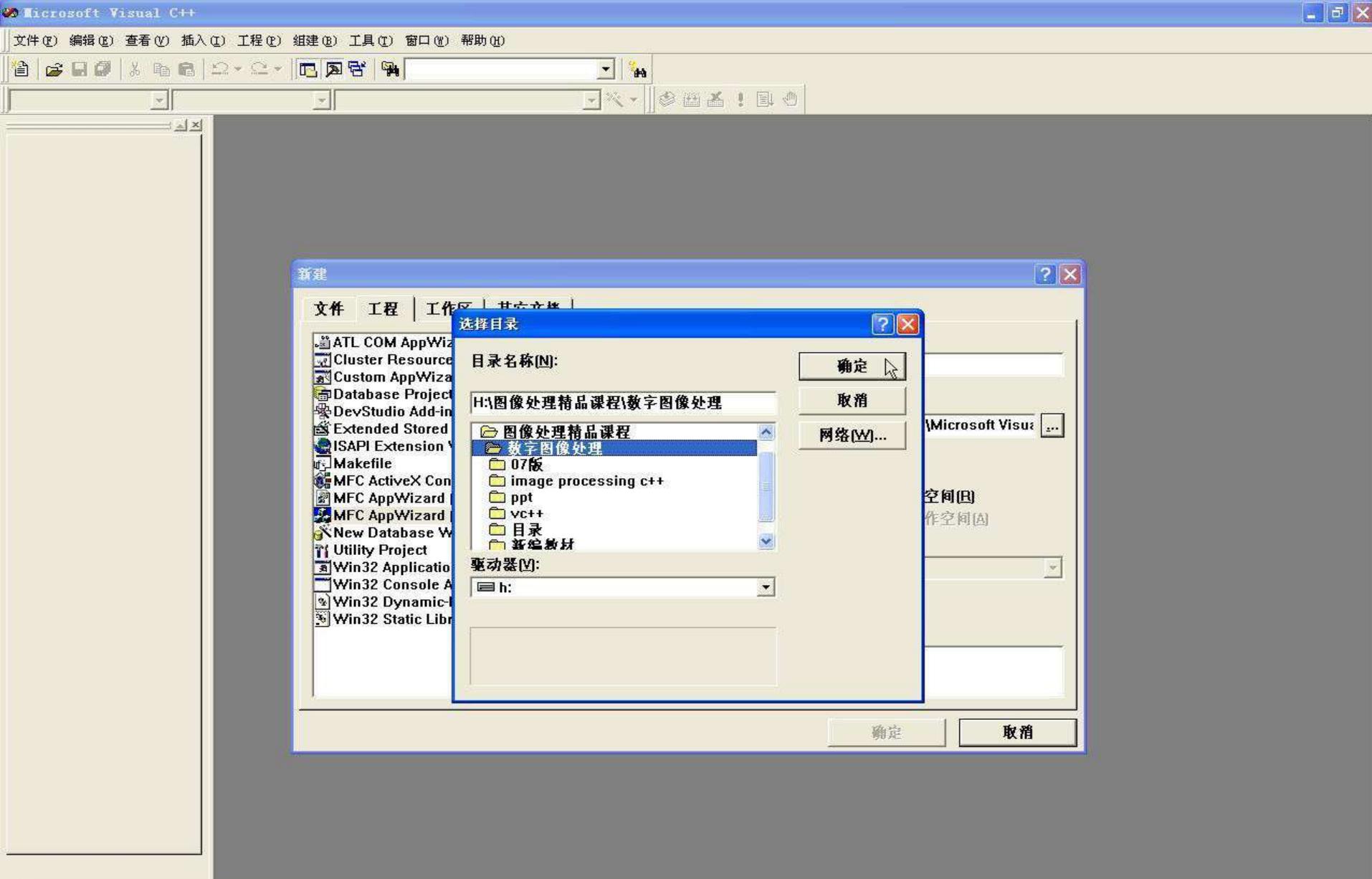
Microsoft Visual...

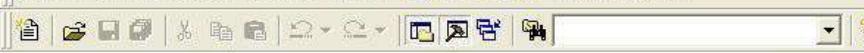
13:09

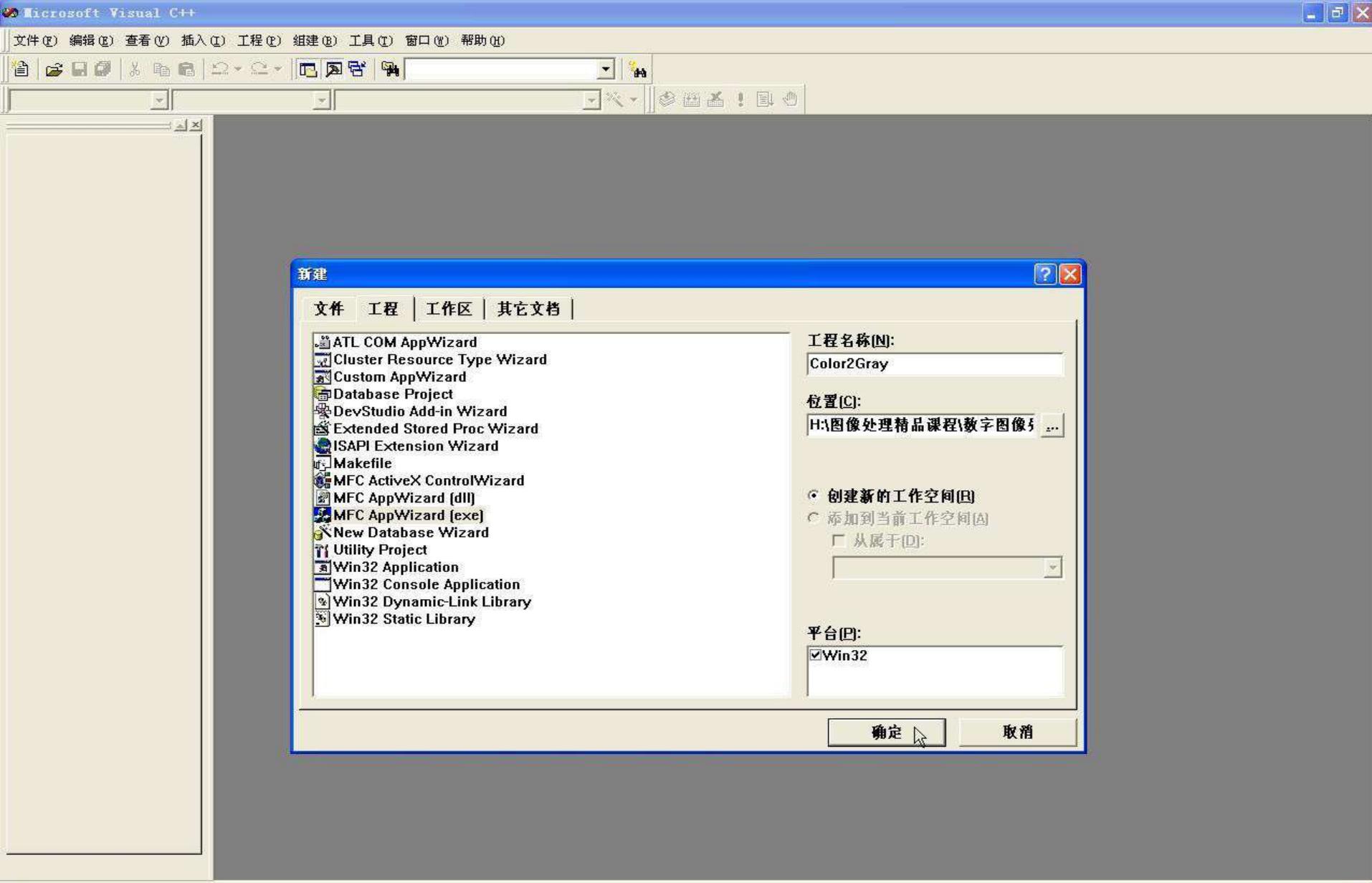




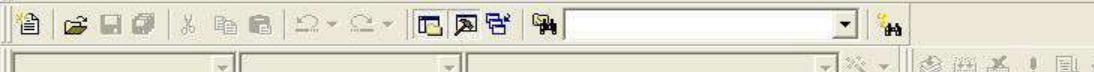














x

新建工程信息



应用程序向导 将会创建一个新的以下规格的工程骨架:

Application type of Color2Gray:

Dialog-Based Application targeting:
Win32

Classes to be created:

Application: CColor2GrayApp in Color2Gray.h and Color2Gray.cpp
Dialog: CColor2GrayDlg in Color2GrayDlg.h and Color2GrayDlg.cpp

Features:

- + About box on system menu
- + 3D Controls
- + Uses shared DLL implementation (MFC42.DLL)
- + ActiveX Controls support enabled
- + Localizable text in:
 中文[中国]

工程目录:

H:\图像处理精品课程\数字图像处理\Color2Gray

确定

取消

x



文件(E) 编辑(E) 查看(V) 插入(I) 工程(E) 组建(B) 布局(L) 工具(T) 窗口(W) 帮助(H)



CColor2GrayDig [All class members] CColor2GrayDig

Color2Gray resource
+ Dialog
+ Icon
+ String Table
+ Version



Cl... Re... Fil...



就绪



Color2Gray - Mic...

精品课程课件

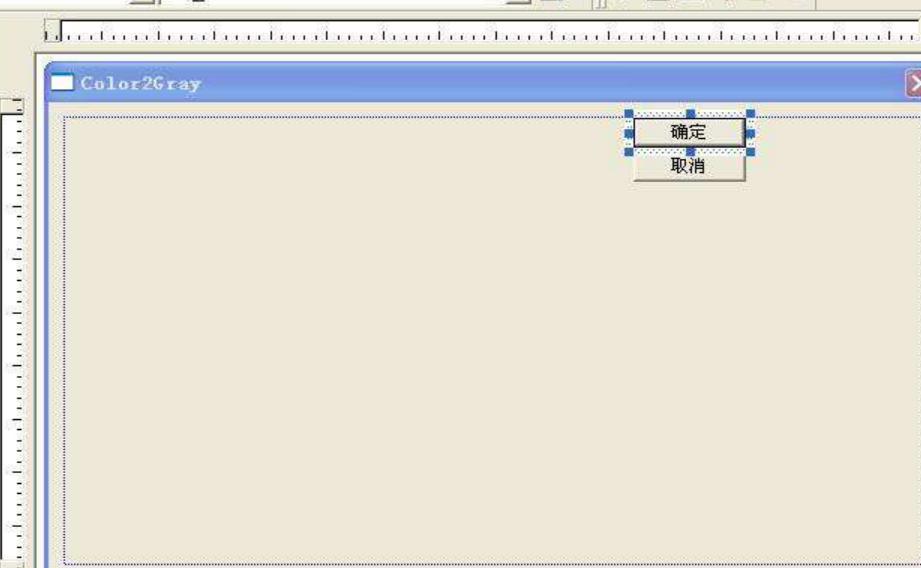
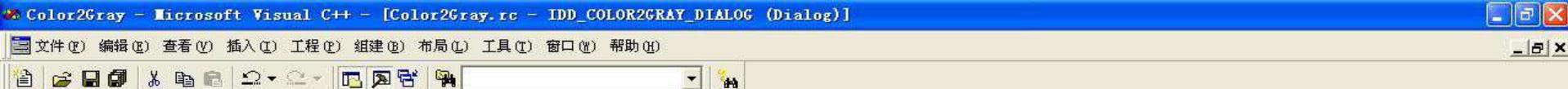
vc(pic)

深蓝

Microsoft PowerP...

未命名 - 画图

14 49



创建一个新的按钮。



Color2Gray - Mic...

精品课程课件

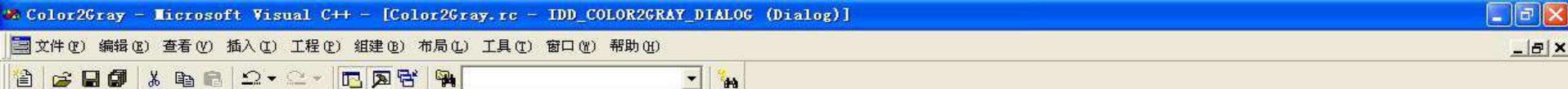
vc(pic)

深蓝

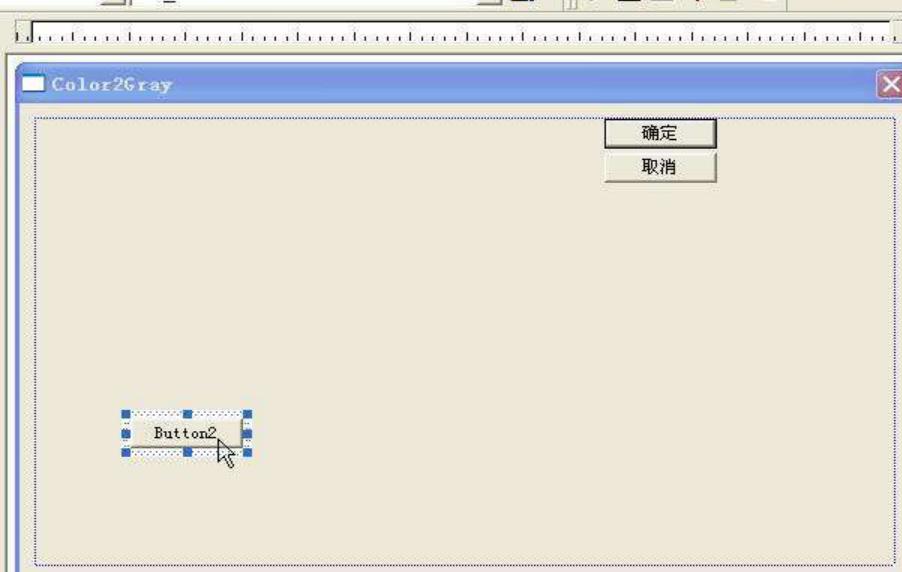
Microsoft PowerP...

远航 - 画图

14:51



CColor2GrayDig IDC_BUTTON2 BN_CLICKED



就绪



Color2Gray - Mic...

精品课程课件

vc(pic)

深蓝

Microsoft PowerP...

Microsoft Word - 画图

49, 148

50 x 14

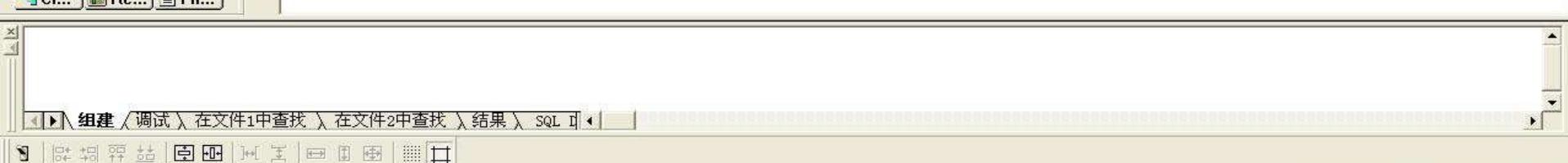
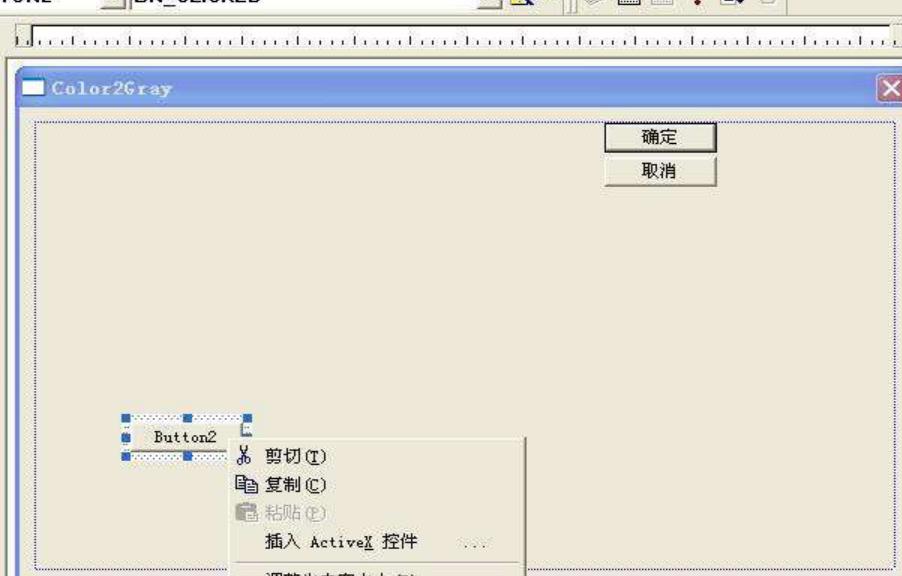
读取

14:53

文件(E) 编辑(E) 查看(V) 插入(I) 工程(E) 组建(B) 布局(L) 工具(T) 窗口(W) 帮助(H)



CCol2GrayDig IDC_BUTTON2 BN_CLICKED



编辑当前被选择的属性



Color2Gray - Mic...

精品课程课件

vc(pic)

深蓝

Microsoft PowerP...

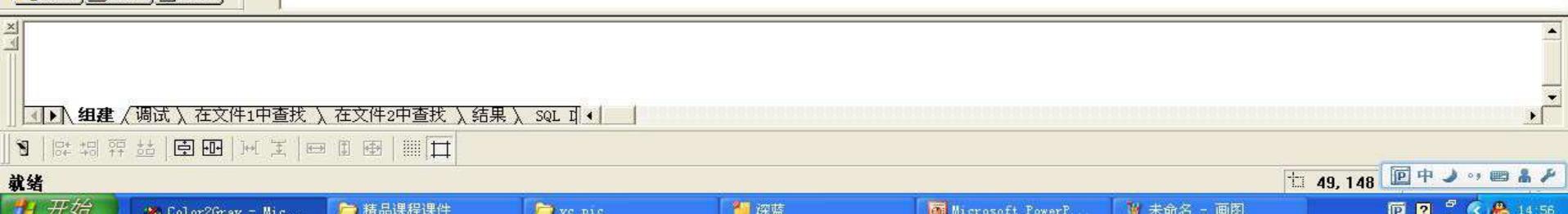
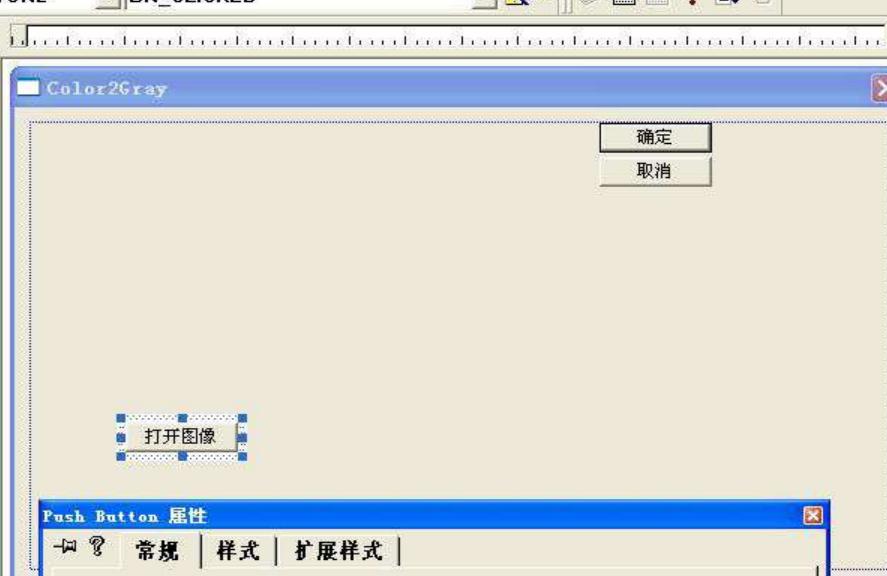
12.JPG - 画图

14:54

文件(E) 编辑(E) 查看(V) 插入(I) 工程(E) 组建(B) 布局(L) 工具(T) 窗口(W) 帮助(H)



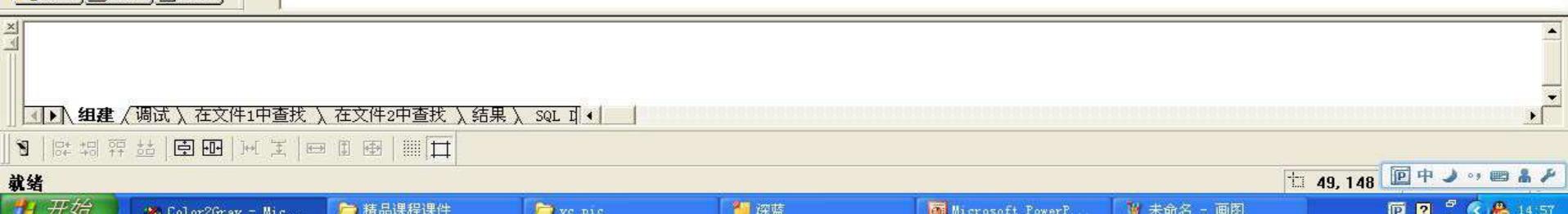
CColor2GrayDig IDC_BUTTON2 BN_CLICKED

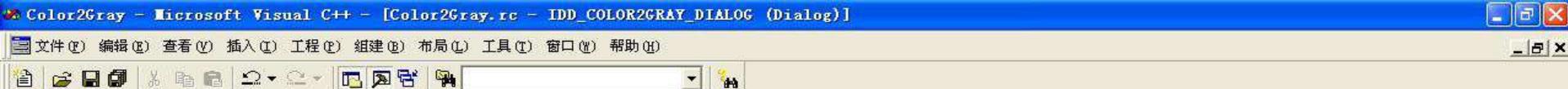


文件(E) 编辑(E) 查看(V) 插入(I) 工程(E) 组建(B) 布局(L) 工具(T) 窗口(W) 帮助(H)



CColor2GrayDig IDC_OPEN BN_CLICKED



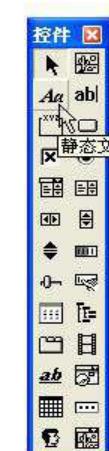


文件(F) 编辑(E) 查看(V) 插入(I) 工程(E) 组建(B) 布局(L) 工具(T) 窗口(W) 帮助(H)



CColor2GrayDig [All class members] CColor2GrayDig

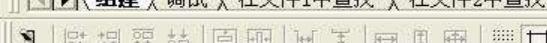
Color2Gray resource
+ Dialog
+ Icon
+ String Table
+ Version



Cl... Re... Fil...

Color2Gray.exe - 0 error(s), 0 warning(s)

组建(Debug) 在文件1中查找 在文件2中查找 结果 SQL D



创建一个新的静态文本控件.



Color2Gray - Mic...

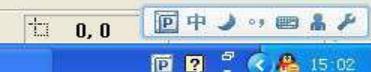
精品课程课件

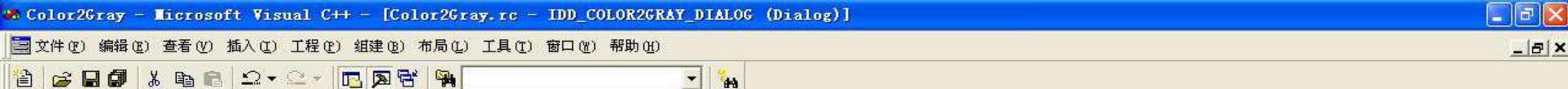
vc(pic)

深蓝

Microsoft PowerP...

15.JPG - 画图



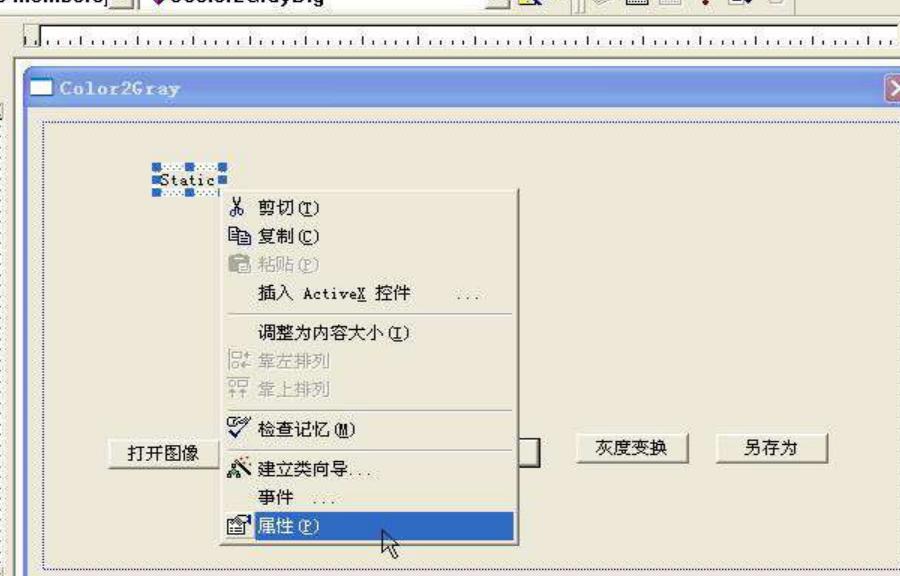


文件(E) 编辑(E) 查看(V) 插入(I) 工程(E) 组建(B) 布局(L) 工具(T) 窗口(W) 帮助(H)



CColor2GrayDig [All class members] CColor2GrayDig

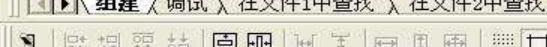
Color2Gray resource
+ Dialog
+ Icon
+ String Table
+ Version



Cl... Re... Fil...

Color2Gray.exe - 0 error(s), 0 warning(s)

组建(Debug) 在文件1中查找 在文件2中查找 结果 SQL D



编辑当前被选择的属性



Color2Gray - Mi...

精品课程课件

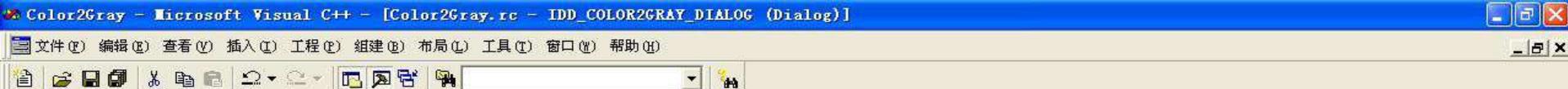
vc(pic)

深蓝

Microsoft PowerP...

16.JPG - 画图

59, 30 中 晴 15:03

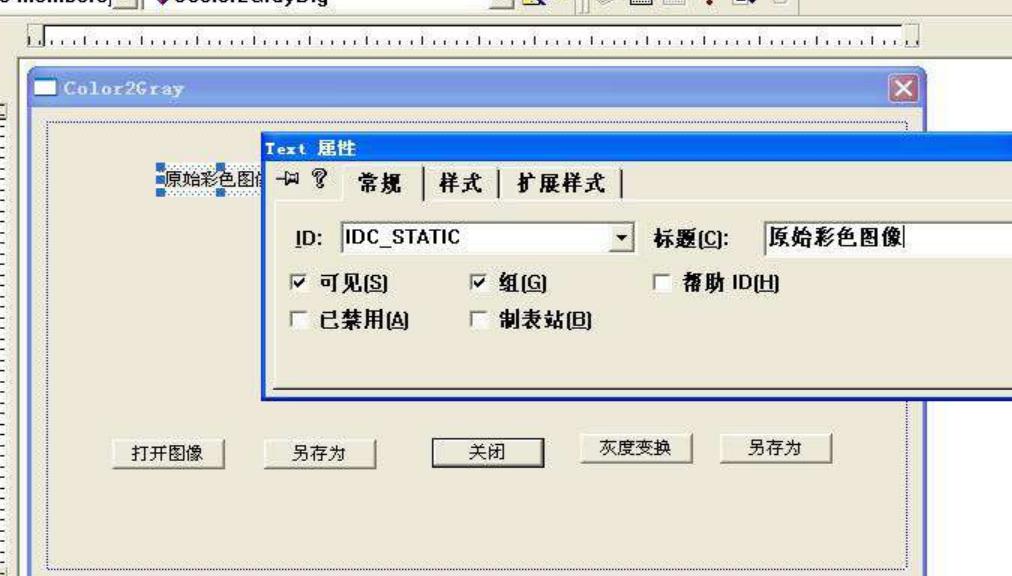


文件(E) 编辑(E) 查看(V) 插入(I) 工程(E) 组建(B) 布局(L) 工具(T) 窗口(W) 帮助(H)



CColor2GrayDig [All class members] CColor2GrayDig

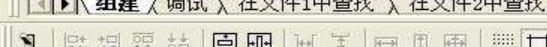
Color2Gray resource
+ Dialog
+ Icon
+ String Table
+ Version



Cl... Re... Fil...

Color2Gray.exe - 0 error(s), 0 warning(s)

组建(Debug) 在文件1中查找 在文件2中查找 结果 SQL D



就绪



Color2Gray - Mic...

精品课程课件

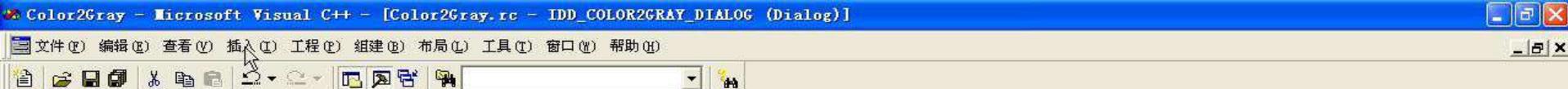
vc(pic)

深蓝

Microsoft PowerP...

17.JPG - 画图





文件(E) 编辑(E) 查看(V) 插入(I) 工程(E) 组建(B) 布局(L) 工具(T) 窗口(W) 帮助(H)

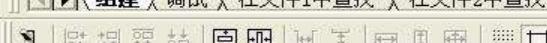


CColor2GrayDig [All class members] CColor2GrayDig



Color2Gray.exe - 0 error(s), 0 warning(s)

组建(Debug) 在文件1中查找 在文件2中查找 结果 SQL



就绪



Color2Gray - Mic...

精品课程课件

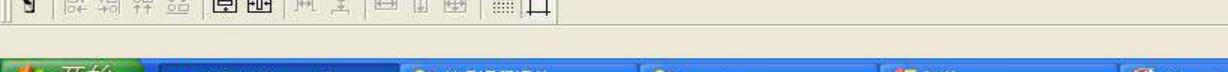
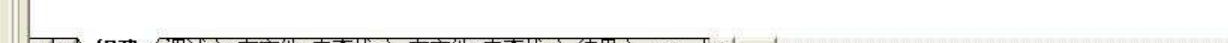
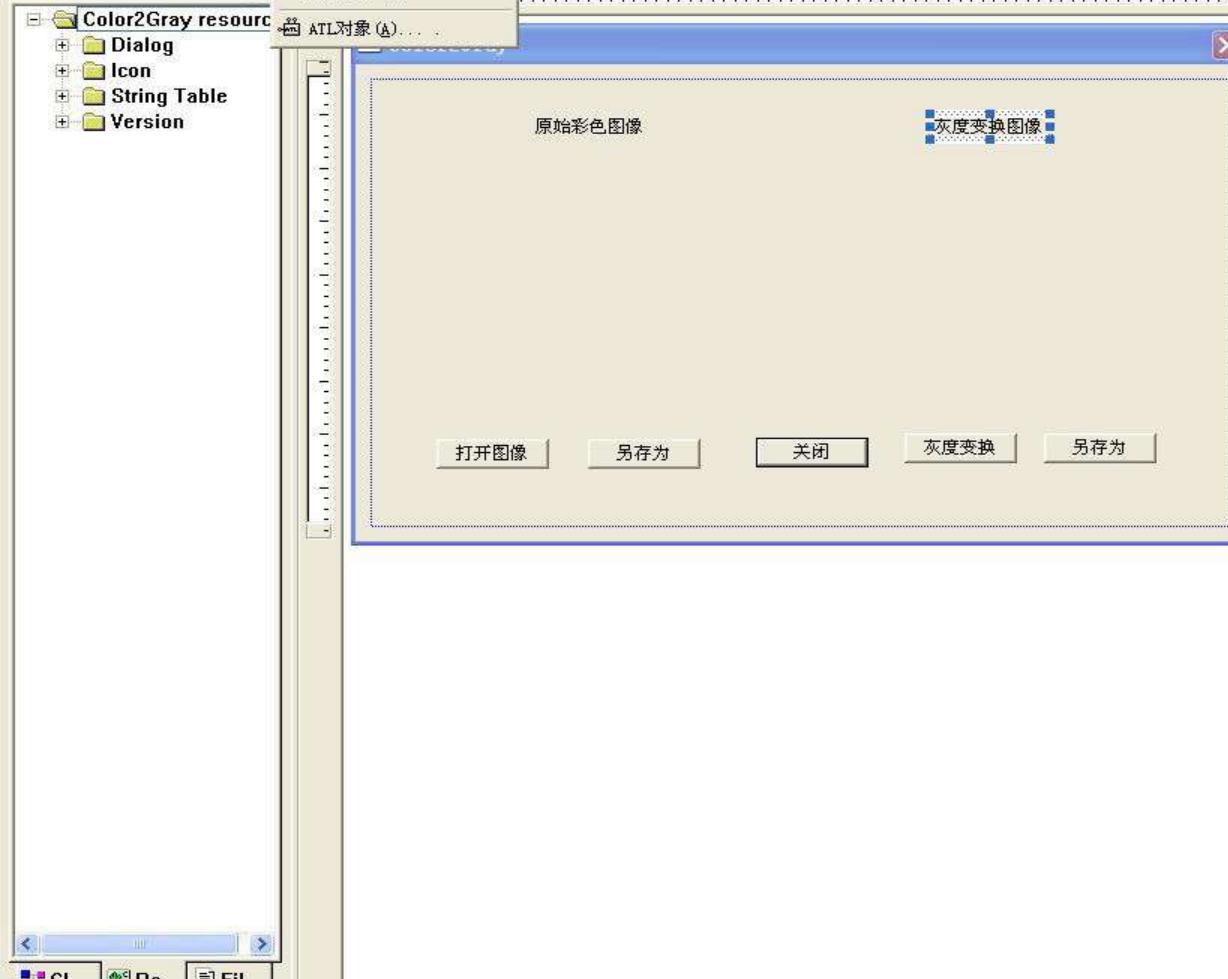
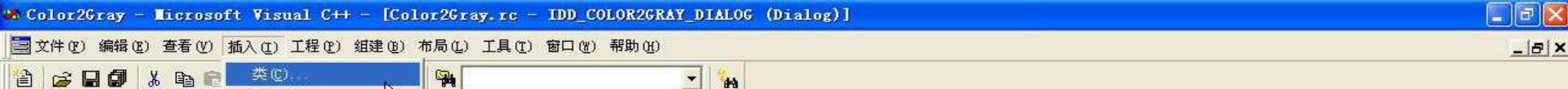
vc(pic)

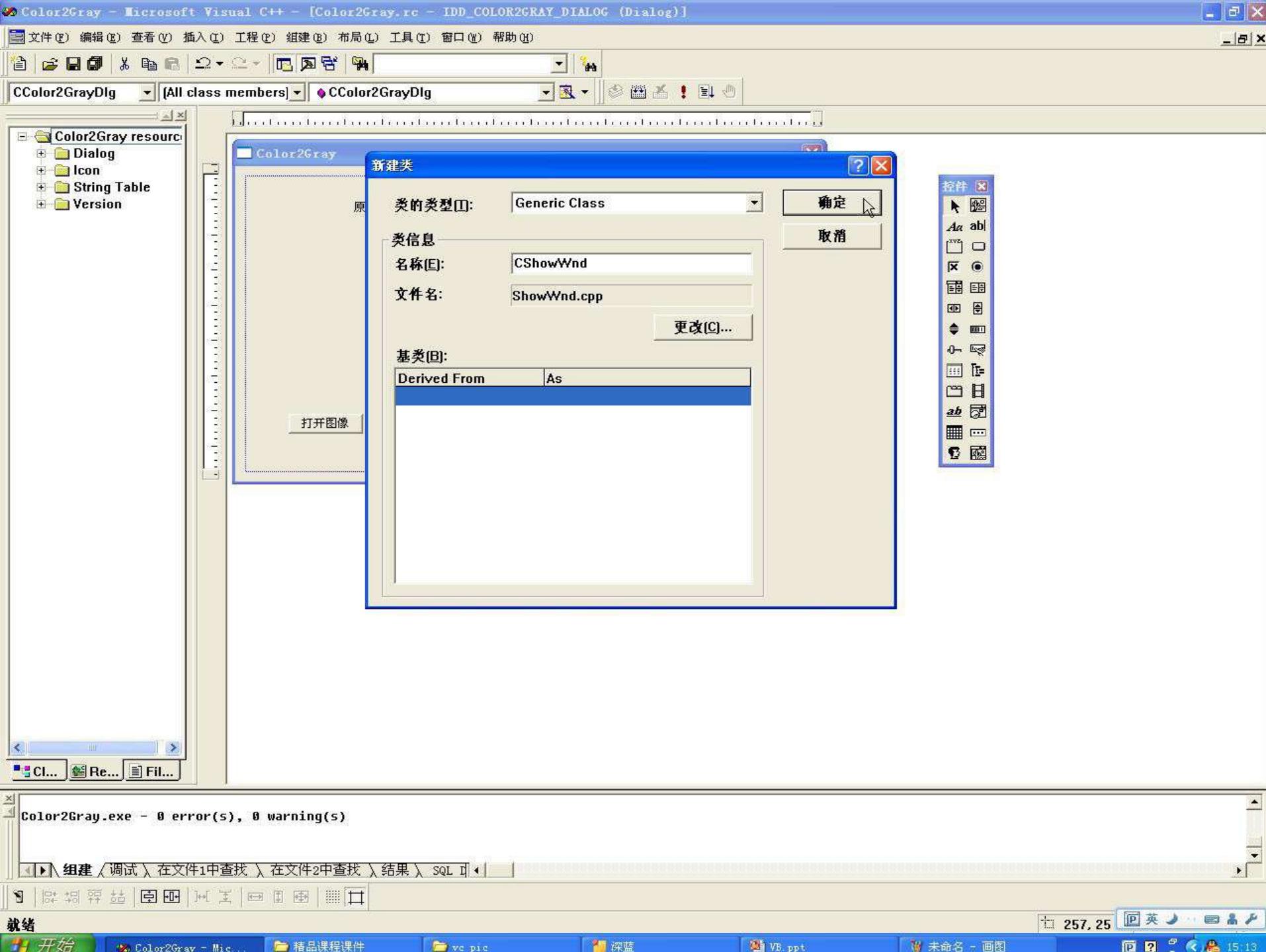
深蓝

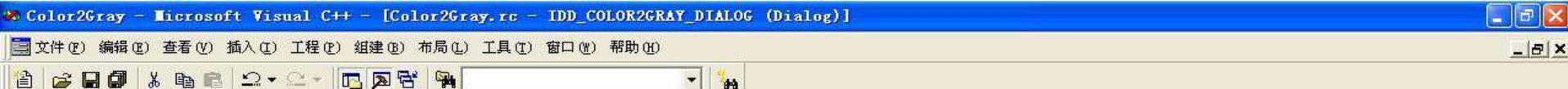
Microsoft PowerP...

未命名 - 画图

257, 25 中 晴 15:07





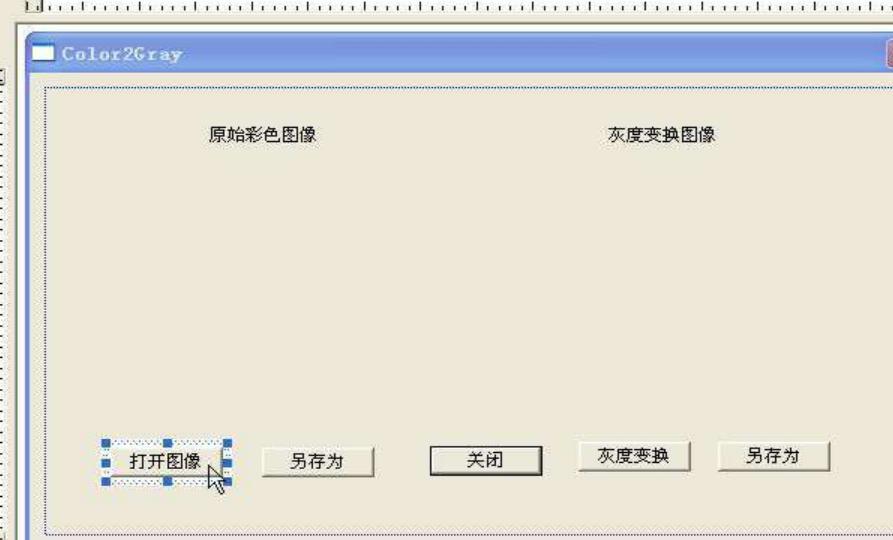


文件(F) 编辑(E) 查看(V) 插入(I) 工程(E) 组建(B) 布局(L) 工具(T) 窗口(W) 帮助(H)



CColor2GrayDlg IDC_OPEN BN_CLICKED

Color2Gray classes
+ CAboutDlg
+ CColor2GrayApp
+ CColor2GrayDlg
+ CShowWnd
+ Globals



Cl... Re... Fil...

Color2Gray.exe - 0 error(s), 0 warning(s)

组建 调试 在文件1中查找 在文件2中查找 结果 SQL D



就绪



Color2Gray - Mic

精品课程课件

vc.pic

深蓝

VB.ppt

未命名 - 画图

36,176



文件(F) 编辑(E) 查看(V) 插入(I) 工程(E) 组建(B) 布局(L) 工具(T) 窗口(W) 帮助(H)



CColor2GrayDig IDC_OPEN BN_CLICKED

Color2Gray classes
+ CABoutDlg
+ CColor2GrayApp
+ CColor2GrayDlg
+ CShowWnd
+ Globals



Cl... Re... Fil...

Color2Gray.exe - 0 error(s), 0 warning(s)

组建 调试 在文件1中查找 在文件2中查找 结果 SQL 工具

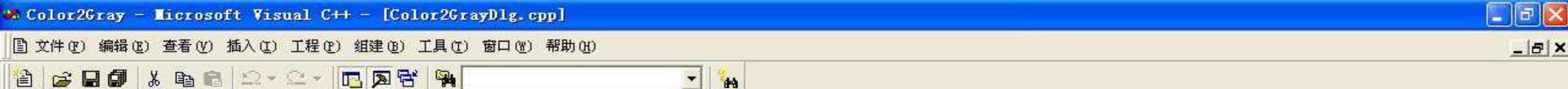
开始 Color2Gray - Mic... 精品课程课件 vc(pic) 深蓝 VB.ppt 未命名 - 画图

就绪

36, 176



15 18



CColor2GrayDlg

[All class members] ▾

OnOpen

```
        }
    else
    {
        CDialog::OnPaint();
    }

// The system calls this to obtain the cursor to display while the user drags
// the minimized window.
HCURSOR CColor2GrayDlg::OnQueryDragIcon()
{
    return (HCURSOR) m_hIcon;
}

void CColor2GrayDlg::OnOpen()
{
    // TODO: Add your control notification handler code here
}
```

Cl... Re... Fil...

Color2Gray.exe - 0 error(s), 0 warning(s)

组建 调试 在文件1中查找 在文件2中查找 结果 SQL

就绪

Color2Gray - Mic... 精品课程课件 vc(pic) 深蓝 VB.ppt 23.JPG - 画图 行 175, 列 1 英

开始

Color2Gray - Microsoft Visual C++ - [Color2GrayDlg.cpp]

文件 (F) 编辑 (E) 查看 (V) 插入 (I) 工程 (P) 组建 (B) 工具 (T) 窗口 (W) 帮助 (H)

CCol2GrayDlg [All class members] OnOpen

Color2Gray classes

- CAboutDlg
- CCol2GrayApp
- CCol2GrayDlg
- CShowWnd
- Globals
- theApp

```
HCURSOR CColor2GrayDlg::OnQueryDragIcon()
{
    return (HCURSOR) m_hIcon;
}

/*
* 函数名称: OnOpen()
* 说 明: 读取指定的BMP格式文件并显示。
*/
void CColor2GrayDlg::OnOpen()
{
    // TODO: Add your control notification handler code here
    //当用户点击“打开图像”按钮时，弹出对话框用于指定BMP格式的文件
    CFile File= false;
    CFileDialog dlg(TRUE, 0, 0, OFN_HIDEREADONLY, "位图文件 (*.bmp)|所有文件 (*.*)|", this);

    //如果用户点击“ok”按钮，则开始尝试读取图像
    if (dlg.DoModal() == IDOK)
    {
        //得到该文件的路径
        FileName = dlg.GetPathName();

        //如果文件打开失败则返回
        if (!File.Open(FileName, CFile::modeRead)) return;

        //若数据指针非空则清空内存
        if (lpBitmap) delete lpBitmap;

        //获取BMP文件的长度(字节)
        nLen = File.GetLength();

        //给该文件数据分配内存空间
        lpBitmap = new BYTE[nLen];

        //将该文件数据读入内存并显示
        File.Read(lpBitmap, nLen);
        LoadBitmap();
        if (lpBitmap) wnd1.ShowImage(nWidth, nHeight, lpBits);
    }
}
```

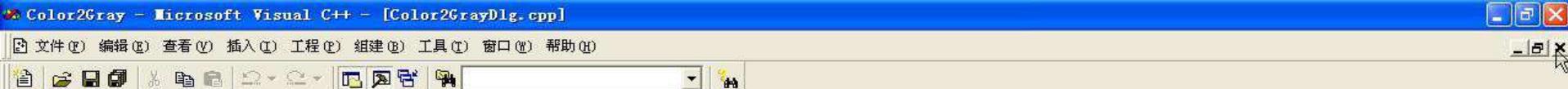
Cl... Re... Fil...

Color2Gray.exe - 0 error(s), 0 warning(s)

组建 调试 在文件1中查找 在文件2中查找 结果 SQL

行 210, 列 3: 英

开始 Color2Gray... Color2Gray... vc(pic) vb(vc matlab) 精品课程课件 深蓝 VB.ppt 24.JPG - 画图 P? 15:27



Color2Gray - Microsoft Visual C++ - [Color2GrayDlg.cpp]

文件(E) 编辑(E) 查看(V) 插入(I) 工程(E) 组建(B) 工具(T) 窗口(W) 帮助(H)

CCol2GrayDlg [All class members] OnOpen

CCol2Gray classes

- CAboutDlg
- CCol2GrayApp
- CCol2GrayDlg
- CShowWnd
- Globals
- theApp

```
CDialog::OnPaint();  
}  
  
// The system calls this to obtain the cursor to display while the user drags  
// the minimized window.  
HCURSOR CCol2GrayDlg::OnQueryDragIcon()  
{  
    return (HCURSOR) m_hIcon;  
}  
  
*****  
* 函数名称: OnOpen()  
* 说 明: 读取指定的BMP格式文件并显示。  
*****  
  
void CCol2GrayDlg::OnOpen()  
{  
    // TODO: Add your control notification handler code here  
    //当用户点击“打开图像”按钮时，弹出对话框用于指定BMP格式的文件  
    CFile File= false;  
    CFileDialog dlg(TRUE, 0, 0, OFN_HIDEREADONLY, "位图文件 (*.bmp)|所有文件 (*.*)|", this);  
  
    //如果用户点击“ok”按钮，则开始尝试读取图像  
    if (dlg.DoModal()==IDOK)  
    {  
        //得到该文件的路径  
        FileName=dlg.GetPathName();  
  
        //如果文件打开失败则返回  
        if (!File.Open(FileName, CFile::modeRead)) return;  
  
        //若数据指针非空则清空内存  
        if (lpBitmap) delete lpBitmap;  
  
        //获取BMP文件的长度(字节)  
        nLen=File.GetLength();  
  
        //给该文件数据分配内存空间  
        lpBitmap=new BYTE[nLen];  
  
        //将该文件数据读入内存并显示  
    }  
}
```

Cl... Re... Fil...

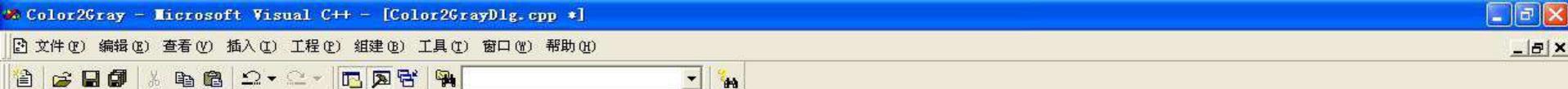


文件(E) 编辑(E) 查看(V) 插入(I) 工程(E) 组建(B) 布局(L) 工具(T) 窗口(W) 帮助(H)



CColor2GrayDig IDC_C2G OnC2g ON_IDC_C2G:BN_CLICK





文件(F) 编辑(E) 查看(V) 插入(I) 工程(P) 组建(B) 工具(T) 窗口(W) 帮助(H)

CCol2GrayDlg [All class members] ♦ OnC2g

```
if (lpBitmap) wnd1.ShowImage(nWidth,nHeight,lpBits);
```

```
/***************************************************************************
* 函数名称: OnC2g()
* 说 明: 将BMP格式的彩色图像转化为256级灰度图像。
* **************************************************************************
```

```
void CColor2GrayDlg::OnC2g()
{
    // TODO: Add your control notification handler code here

    //如果图像数据的指针为空则返回
    if (lpBitmap==0) return;

    //定义图像的横、纵坐标(循环变量)以及当前像素点的位置
    int x,y,p;
    BYTE Point;

    //循环过程, 对rgb图像中的每个像素进行灰度变换
    for(y=0;y<nHeight;y++)
    {
        for(x=0;x<nWidth;x++)
        {
            p=x*3+y*nByteWidth;
            Point=(BYTE)(0.31*(float)lpBits[p+2]+0.59*(float)lpBits[p+1]+0.1*(float)lpBits[p]);
            lpBits[p+2]=Point;
            lpBits[p+1]=Point;
            lpBits[p]=Point;
        }
        //显示变换后的灰度图像
        wnd2.ShowImage(nWidth,nHeight,lpBits);
    }
}
```

Cl... Re... Fil...

Color2Gray.exe - 0 error(s), 0 warning(s)

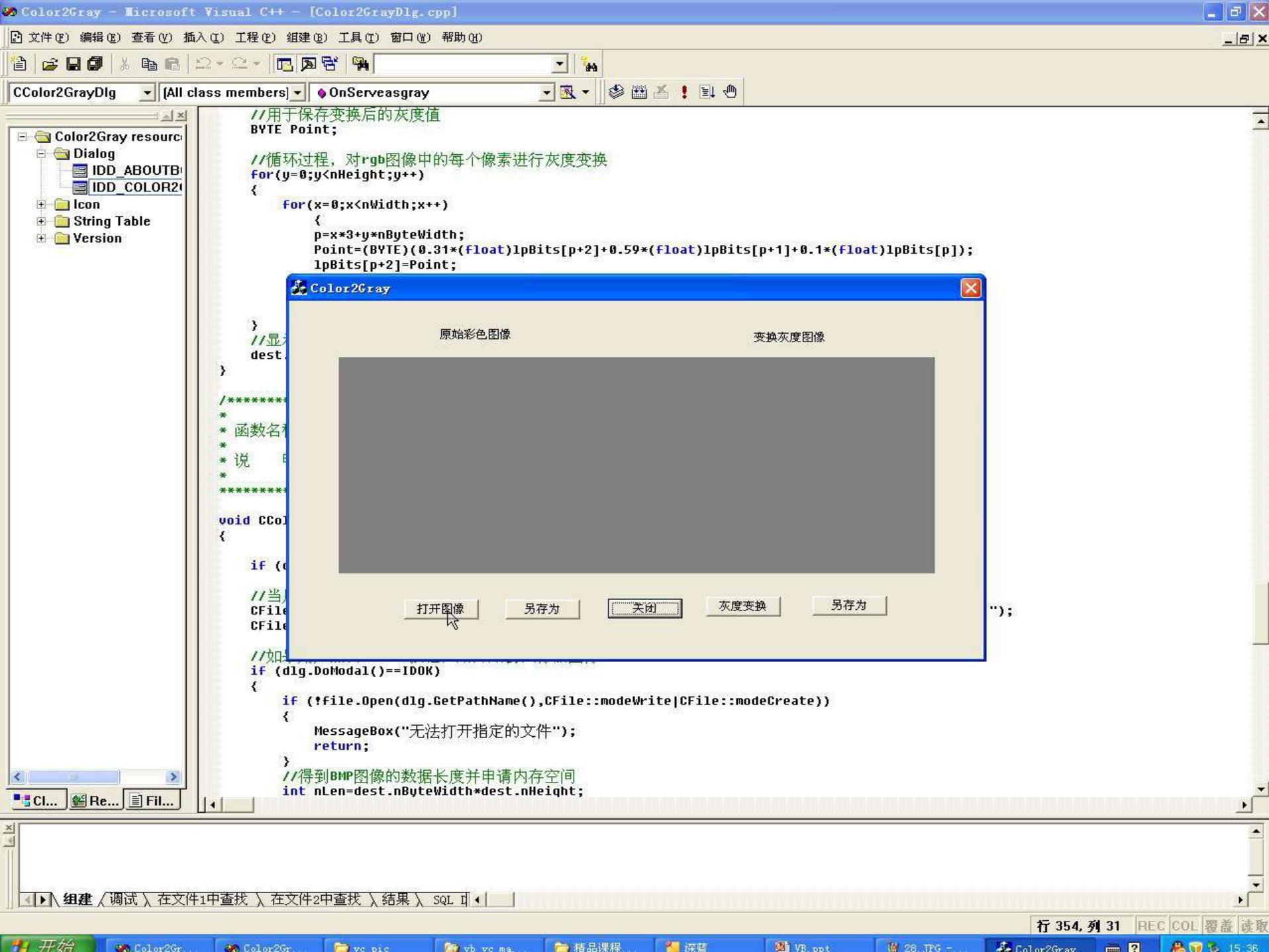
组建 调试 在文件1中查找 在文件2中查找 结果 SQL

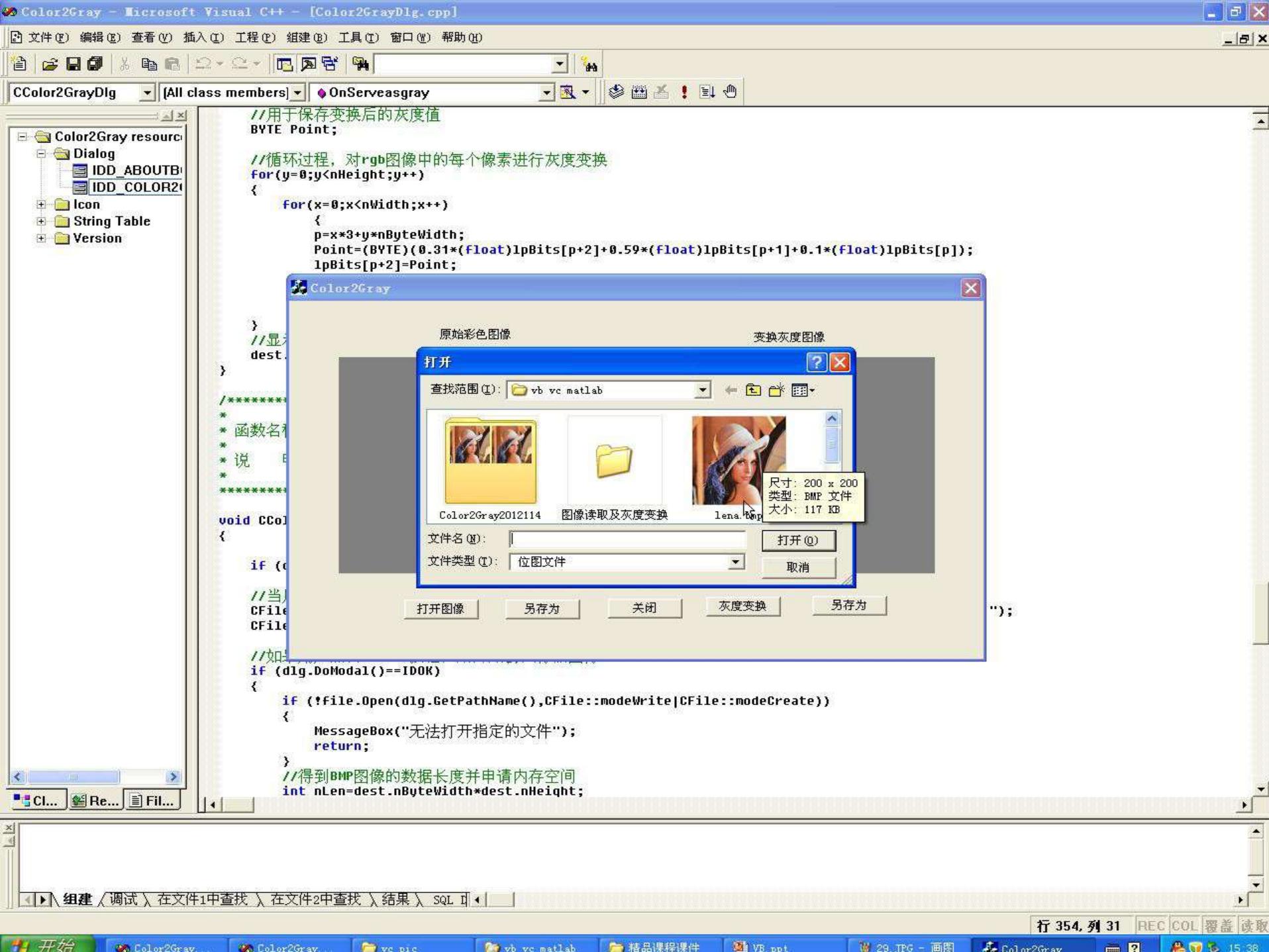
就绪

Color2Gray... Color2Gray... vc(pic) vb(vc matlab) 精品课程课件 深蓝 VB.ppt 未命名 - 画图

行 240, 列 1 英

15:36





Color2Gray - Microsoft Visual C++ - [Color2GrayDlg.cpp]

文件 (F) 编辑 (E) 查看 (V) 插入 (I) 工程 (P) 组建 (B) 工具 (T) 窗口 (W) 帮助 (H)

CCol2GrayDig [All class members] ♦ OnServeasgray

//用于保存变换后的灰度值
BYTE Point;

//循环过程, 对rgb图像中的每个像素进行灰度变换
for(y=0;y<nHeight;y++)
{
 for(x=0;x<nWidth;x++)
 {
 p=x*3+y*nByteWidth;
 Point=(BYTE)(0.31*(float)lpBits[p+2]+0.59*(float)lpBits[p+1]+0.1*(float)lpBits[p]);
 lpBits[p+2]=Point;

}
 //显示
 dest.

* 函数名和
* 说

void CCol2GrayDig::OnServeasgray()

if (0)

//当
CFile file("D:\\Color2Gray\\Color2Gray.bmp", CFile::modeRead);
CFile file("D:\\Color2Gray\\Color2Gray.bmp", CFile::modeWrite);
//如
if (dlg.DoModal()==IDOK)
{
 if (!file.Open(dlg.GetPathName(),CFile::modeWrite|CFile::modeCreate))
 {
 MessageBox("无法打开指定的文件");
 return;
 }
 //得到BMP图像的数据长度并申请内存空间
 int nLen=dest.nByteWidth*dest.nHeight;

Color2Gray

原始彩色图像 变换灰度图像

打开图像 另存为 关闭 灰度变换 另存为

Color2Gray - Microsoft Visual C++ - [Color2GrayDlg.cpp]

文件 (F) 编辑 (E) 查看 (V) 插入 (I) 工程 (P) 组建 (B) 工具 (T) 窗口 (W) 帮助 (H)

CColor2GrayDig [All class members] ♦ OnServeasgray

```
//用于保存变换后的灰度值  
BYTE Point;  
  
//循环过程, 对rgb图像中的每个像素进行灰度变换  
for(y=0;y<nHeight;y++)  
{  
    for(x=0;x<nWidth;x++)  
    {  
        p=x*3+y*nByteWidth;  
        Point=(BYTE)(0.31*(float)lpBits[p+2]+0.59*(float)lpBits[p+1]+0.1*(float)lpBits[p]);  
        lpBits[p+2]=Point;  
    }  
}  
//显示  
dest.  
*****  
* 函数名  
* 说  
*  
*****  
  
void CColor2GrayDlg::OnPaint()  
{  
    if (IsIconic())  
    {  
        //当  
        CFile file;  
        CFile dest;  
        //如  
        if (dlg.DoModal()==IDOK)  
        {  
            if (!file.Open(dlg.GetPathName(),CFile::modeWrite|CFile::modeCreate))  
            {  
                MessageBox("无法打开指定的文件");  
                return;  
            }  
            //得到BMP图像的数据长度并申请内存空间  
            int nLen=dest.nByteWidth*dest.nHeight;
```

Color2Gray

原始彩色图像 变换灰度图像





数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第1章 数字图像处理的基本知识

第4讲 数字图像处理系统简介及数字图像处理应用



第1章 数字图像处理的基本知识

1.3 数字图像处理系统简介

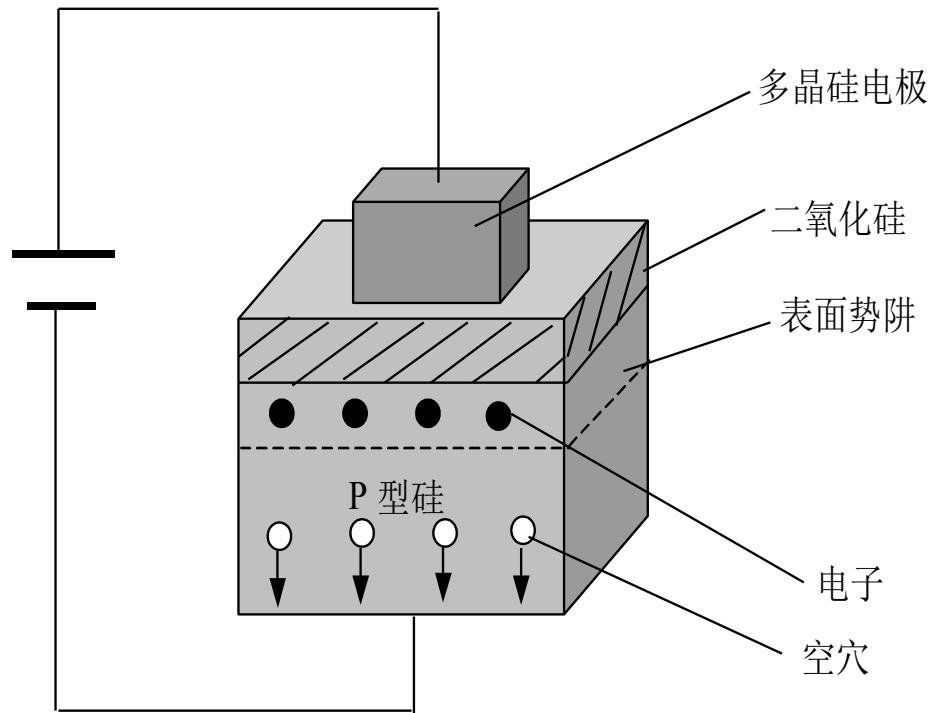
1.4 数字图像处理应用



第1章 数字图像处理的基本知识

1.3 数字图像处理系统简介

CCD 摄像器件的工作原理：CCD 器件是一种 MOS 集成电路，在 P 型硅基片上用氧化法生成二氧化硅薄层作为绝缘层，在绝缘层上用多晶硅做成电极，构成一个具有 MOS 结构的电容器，在电极上加正电压，在 P 型硅基片的电极下面就会形成可存储电荷的势阱。光从电极旁边照射时，就会产生电子—空穴对，少数载流子就会存储在势阱中，形成所谓电荷包，且电荷数量与光照强弱成比例。



单个 MOS 电容器的构造



第1章 数字图像处理的基本知识

1.4 数字图像处理应用

➤ 图像处理技术的发展

1839年——照相术发明

1865年——法国巴黎和里昂之间文字传真

1893年——发明电影

1914年——首张新闻传真照片出现在媒体

1925年——发明电视

1960年——NTSC 彩色电视开播

1964年——航天探测器徘徊者7号发回地面几千张月球表面照片（美国喷气推进实验室JPL）。由计算机成功绘制月球表面地图。



第1章 数字图像处理的基本知识

1.4 数字图像处理应用

1972年——发明了X射线断层摄像装置，简称CT，用于头颅诊断。

1975年——研制出用于全身的CT装置，1979年该无损诊断技术获得诺贝尔奖。

1980年——CCITT(国际电报电话咨询委员会)制定了三类传真机和公用电话网传输静止图像的国际标准。

1984年——CCITT提出的第一个实用化的、适用于会议电视和可视电话要求的视频压缩编码国际标准，用于在ISDN网络上上传输实际活动图像。

之后，数字图像处理应用得到迅速发展，图像成为人们生活不可或缺的重要组成部分。CCD，数码相机，……。

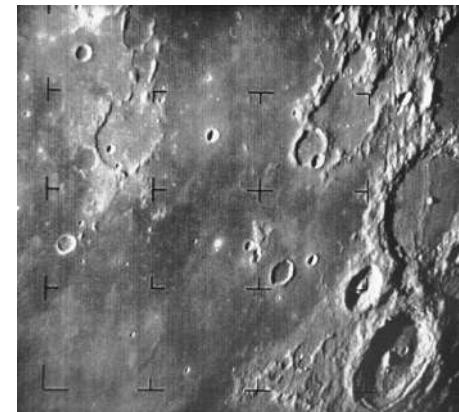


第1章 数字图像处理的基本知识

1.4 数字图像处理应用



1921 年的打印图像



1964 年美国航天器发
回的第一张月球照片



CT 图像





第1章 数字图像处理的基本知识

1.4 数字图像处理应用

➤ 应用领域：

物理、化学：结晶分析、谱分析；

生物、医学：细胞分析、染色体分类、血球分类、X射线照片分析、CT
；

环境保护：水质及大气污染调查；

地质：资源勘探、地图绘制、GIS；

农林：植被分布调查、农作物估产；

海洋：鱼群探查、海洋污染监测；

水利：河流分布、水利及水害调查；

气象：云图分析等；

通信：传真、电视、多媒体通信；

工业、交通：工业探伤、铁路选线、机器人、产品质量监测；

经济：电子商务、身份认证、防伪；

军事：军事侦察、导弹制导、电子沙盘、军事训练等；

其它：指纹识别等。

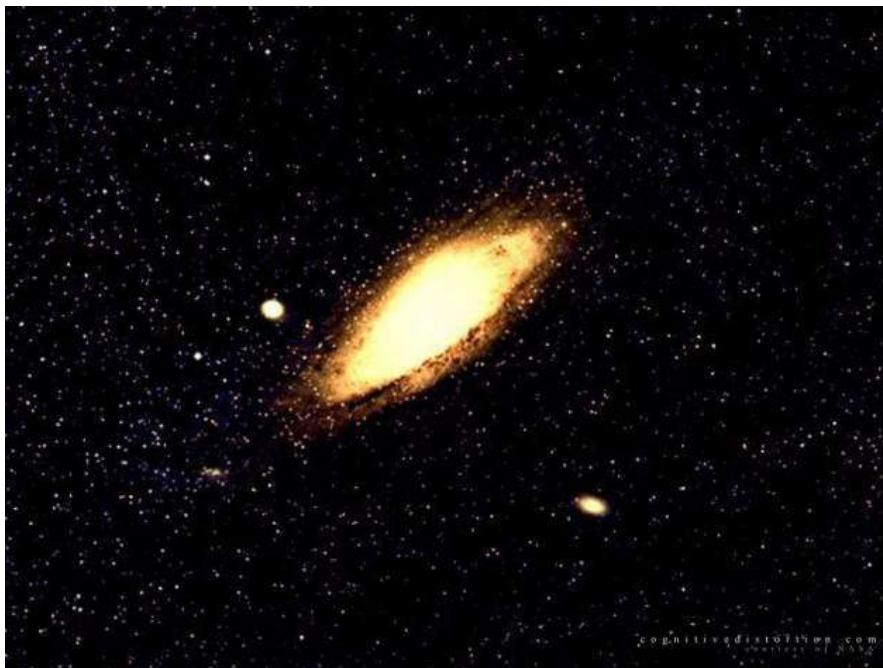




第1章 数字图像处理的基本知识

1.4 数字图像处理应用

图像处理最先应用于空间探索

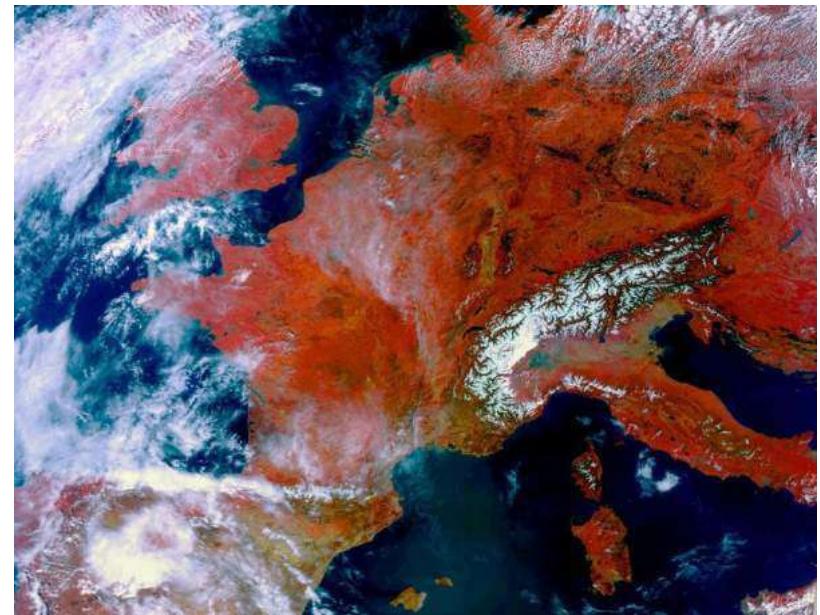




第1章 数字图像处理的基本知识

1.4 数字图像处理应用

地球资源勘探



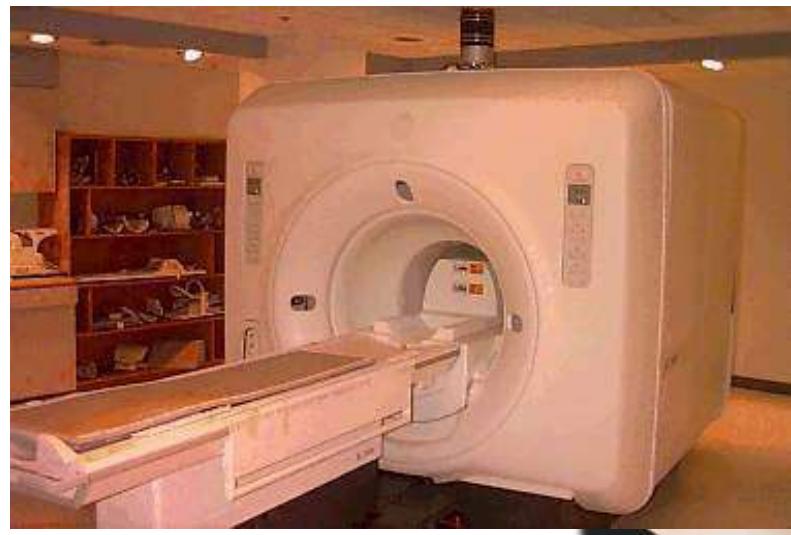
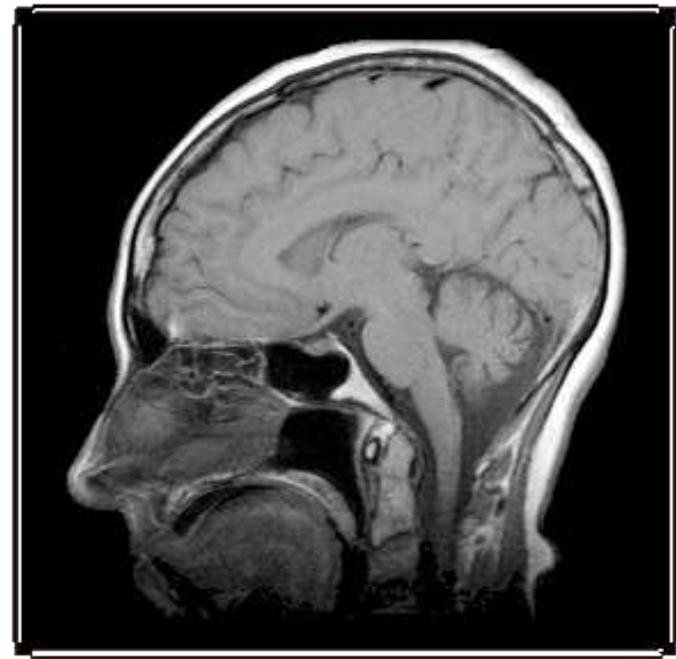
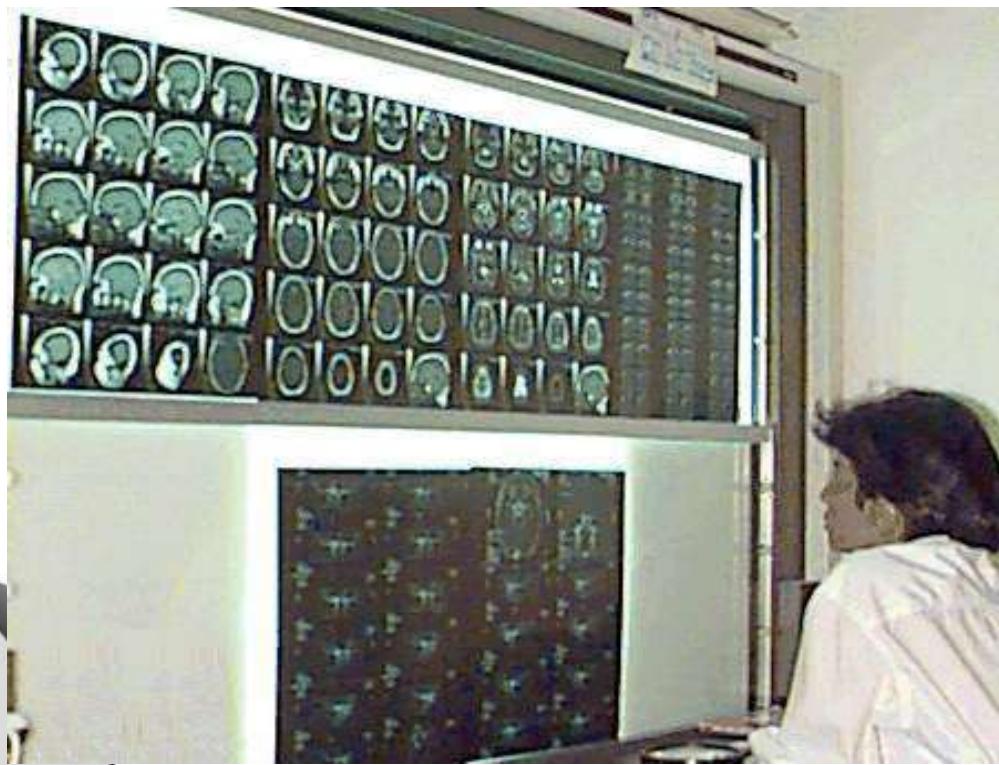
遥感图像



第1章 数字图像处理的基本知识

1.4 数字图像处理应用

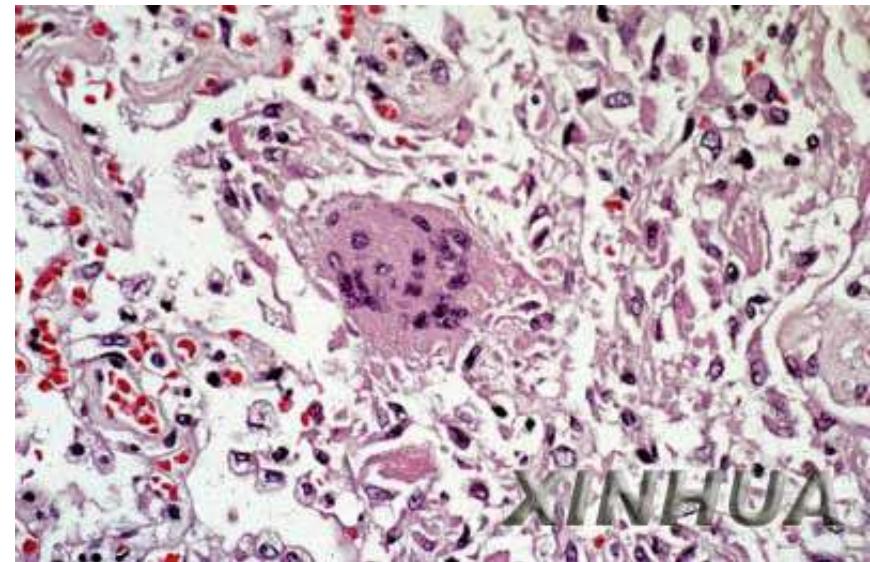
MRI





第1章 数字图像处理的基本知识

1.4 数字图像处理应用



显微图像

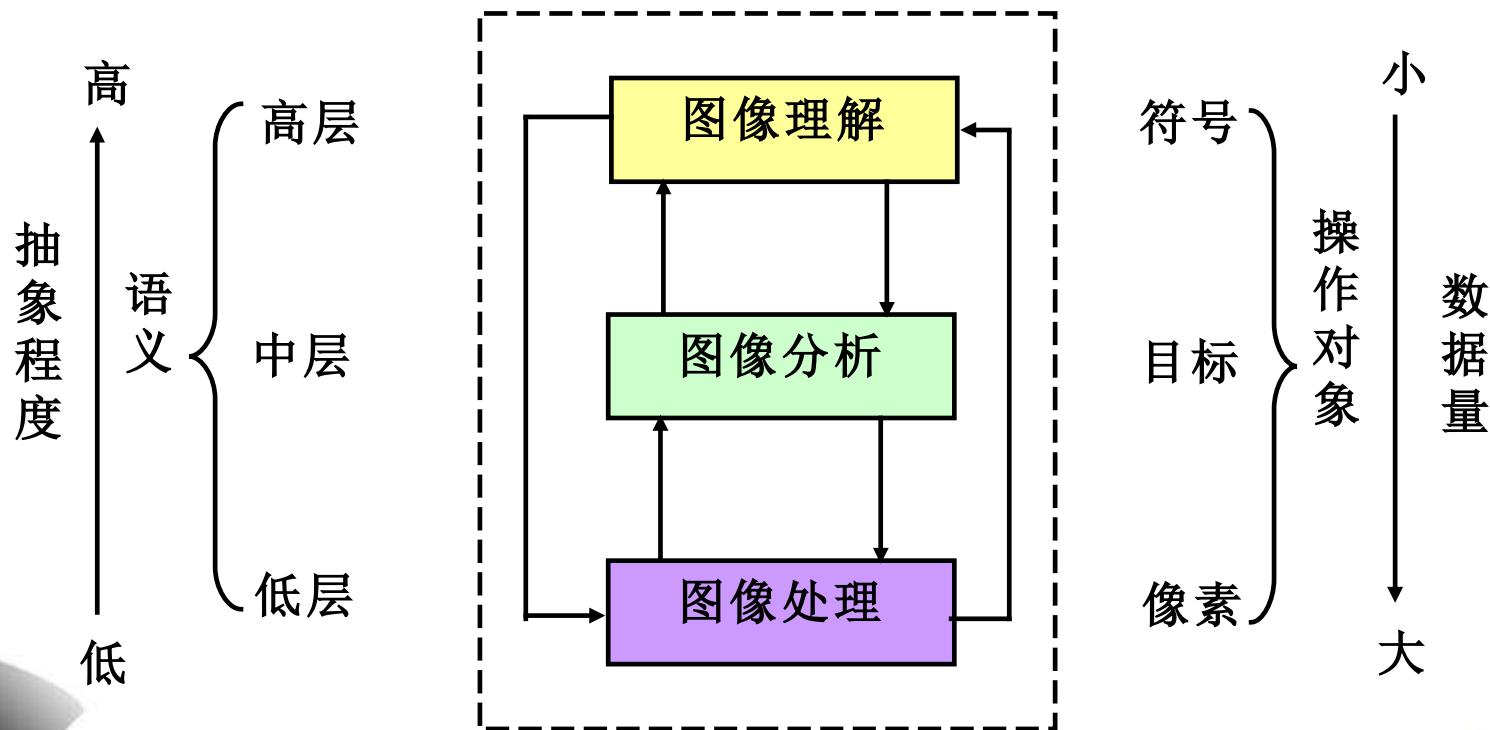


第1章 数字图像处理的基本知识

1.4 数字图像处理应用

➤ 图像处理与图像技术

图像技术是各种与图像有关的技术总称，根据抽象程度和研究方法的不同，可以分为图像处理、图像分析和图像理解三个层次。

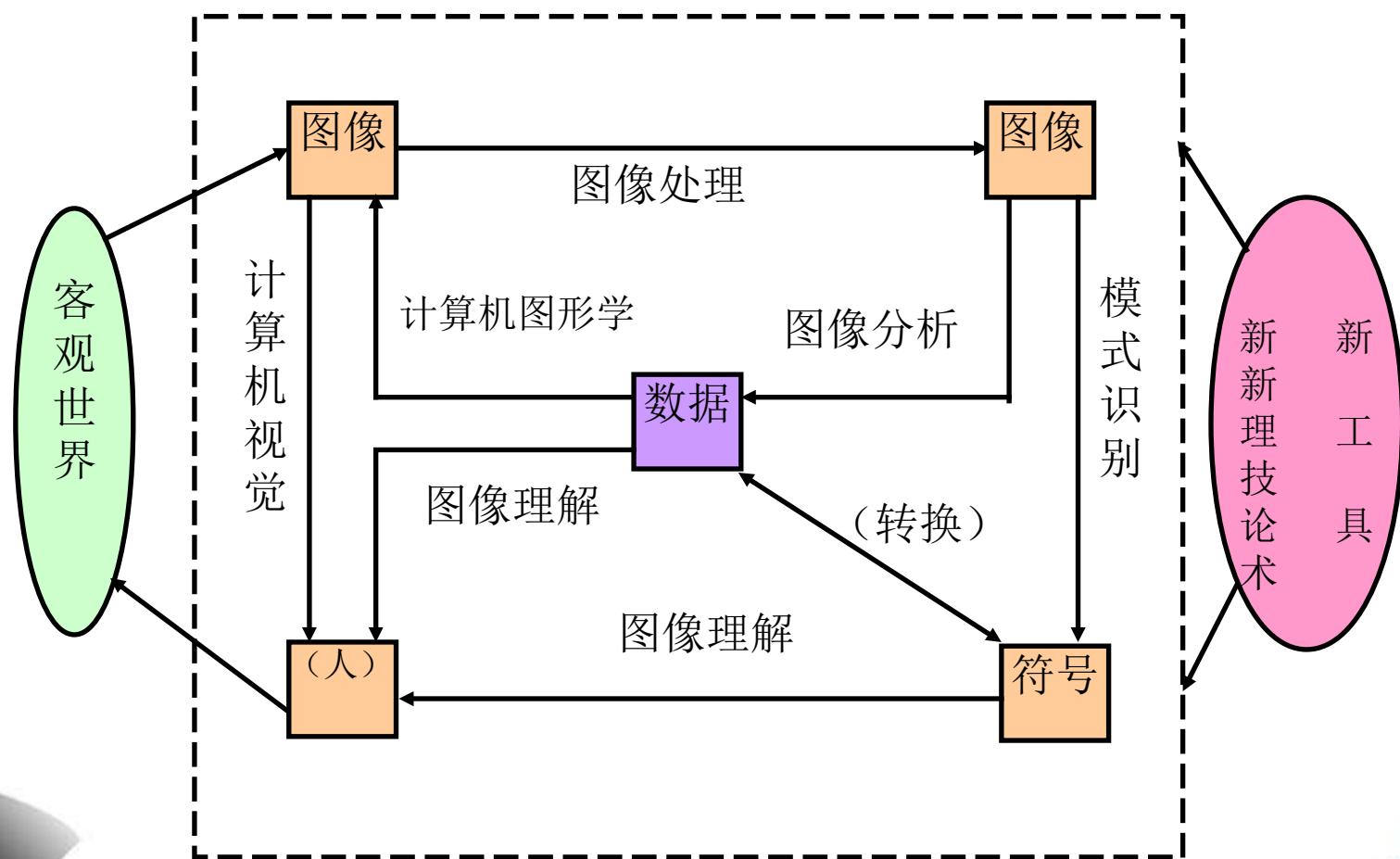




第1章 数字图像处理的基本知识

1.4 数字图像处理应用

➤ 图像处理与相关学科关系





第1章 数字图像处理的基本知识

小结：本章从图像处理的基本概念、图像的表示、图像处理领域的发展与应用、图像处理与其他相关学科的关系、图像处理系统、典型的成像方法、图像显示技术等方面作了简要介绍，使读者对图像处理有了一个概括性的了解。对数字图像常用的几种存储格式进行了介绍，特别是采用目前常用的几种编程软件（VB、VC、Matlab）如何实现图像的读写和存储进行了详细的介绍，将对后续内容的学习中为理论与编程实现的结合奠定了基础。



本课程主要内容

第1章 数字图像处理的基本知识

第2章 图像处理中的常用数学变换

第3章 图像增强

第4章 图像复原

第5章 图像压缩编码

第6章 图像分割

第7章 图像描述

第8章 数字图像技术应用系统设计



参考书

曹茂永	数字图像处理	北京大学出版社
贾永红	数字图像处理	武汉大学出版社
阮秋琦	数字图像处理学	电子工业出版社
杨淑莹	VC++ 图像处理程序设计	清华大学出版社
章毓晋	图象工程：图象处理和分析	清华大学出版社
容观澳	计算机图象处理	清华大学出版社
冈萨雷斯（美）	数字图像处理（英文版）	电子工业出版社
卡斯特曼（美）	数字图像处理（英文版）	电子工业出版社



课程信息

1. 课程性质

专业技术课：电子信息类(电子信息工程、电子科学与技术、通信工程、微电子科学与工程、光电信息科学与工程和信息工程)，遥感测绘类(测绘工程、遥感科学与技术)，计算机类(计算机科学与技术、软件工程、网络工程、信息安全、物联网工程和数字媒体技术)

➤ 学时：54学时（其中实验10学时）

➤ 学分：3学分

2. 考核方式

总分100分，考试70分，平时成绩10分(作业、课堂听课情况等)，实验20分。为了鼓励学生自主学习和参加创新活动，用图像处理技术参加各类创新竞赛获奖、参加老师科研课题成绩突出、自主学习效果明显等，附加分10分，计入总成绩。但满分100分封顶。



课程信息

3. 课程资源

数字图像处理国家级精品课程 (<http://jpkc.sust.edu.cn/sztxcl/>) , 网上学习和答疑。

实验室开放 (地点: J11-318)

4. 联系方式

电子信息类: 曹茂永

[E-mail : my-cao@263.net](mailto:my-cao@263.net)

计算机类: 郑永果

[E-mail : zhengyg206@163.com](mailto:zhengyg206@163.com)

遥感测绘类: 江 涛

[E-mail : tjian@126.com](mailto:tjian@126.com)



国家级精品资源共享课

数字图像处理

(Digital Image Processing)

山东科技大学
江涛教授



第二章 常用的数学变换

2.1 引言

2.2 空域变换

2.3 频率域变换

2.4 离散余弦变换

2.5 K-L 变换

2.6 其他正交变换



第二章 常用的数学变换

2.1 引言

图像的数学变换的特点在于其有精确的数学背景，是许多图像处理技术的基础。在这些变换中，一种是在空间域上进行的，这些变换根据处理操作的特点，可以分为图像的代数运算和几何运算，它们都是利用对输入图像进行加工而得到输出图像。另一种重要的数学变换则是将原定义在图像空间的图像以某种形式转换到另外一些空间，并利用输入图像在这些空间的特有性质有效而快速地对图像进行处理和分析。最典型的变换有离散傅立叶变换，它把空域中的图像信号看作二维时间序列，将其变换到频率域来分析图像的频谱特性。



第二章 常用的数学变换

2.1 引言

空域变换有如加法、减法等的代数变换，也有如旋转、拉伸等的几何变换；频域变换除了傅立叶变换外，常用的非空域的变换还有离散余弦变换、PCA 变换等等。无论是在空域中的数学变换还是频域中的数学变换，它们在图像分析、滤波、增强、压缩等处理中都有着非常典型而重要的应用。



第二章 常用的数学变换

2.1 引言

□ 图像的像素级运算

◆ 点运算

➤ 线性点运算、非线性点运算

◆ 代数运算

➤ 加法、减法、乘法、除法

◆ 逻辑运算

➤ 求反、异或、或、与

□ 图像的空域变换

◆ 几何变换

◆ 非几何变换

➤ 直方图变换





第二章 常用的数学变换

2.1 引言

➤ 线性点运算

$$I_{out}(x, y) = a \cdot I_{in}(x, y) + b$$

- 当 $a = 1, b = 0$ 时 恒等
- 当 $a < 0$ 时 黑白翻转
- 当 $|a| > 1$ 时 增加对比度
- 当 $|a| < 1$ 时 减小对比度
- 当 $b > 0$ 时 增加亮度
- 当 $b < 0$ 时 减小亮度



第二章 常用的数学变换

2.1 引言

➤ 非线性点运算

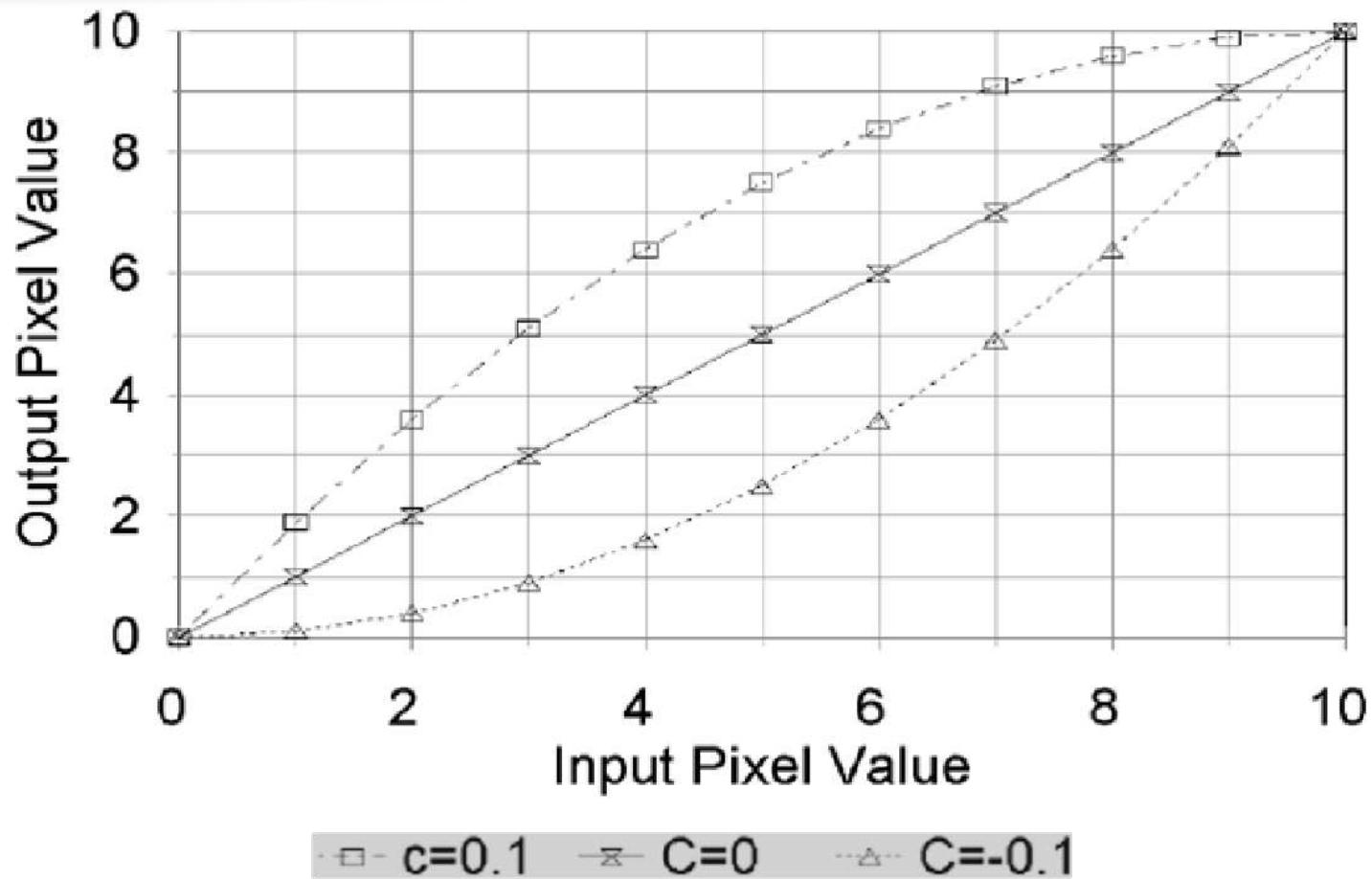
$$f[I(x, y)] = I(x, y) + C \cdot I(x, y) \cdot [I(x, y)_m - I(x, y)]$$

- 当 $C < 0$ 时，增强中间部分亮度
- 当 $C > 0$ 时，奸笑中间部分亮度



第二章 常用的数学变换

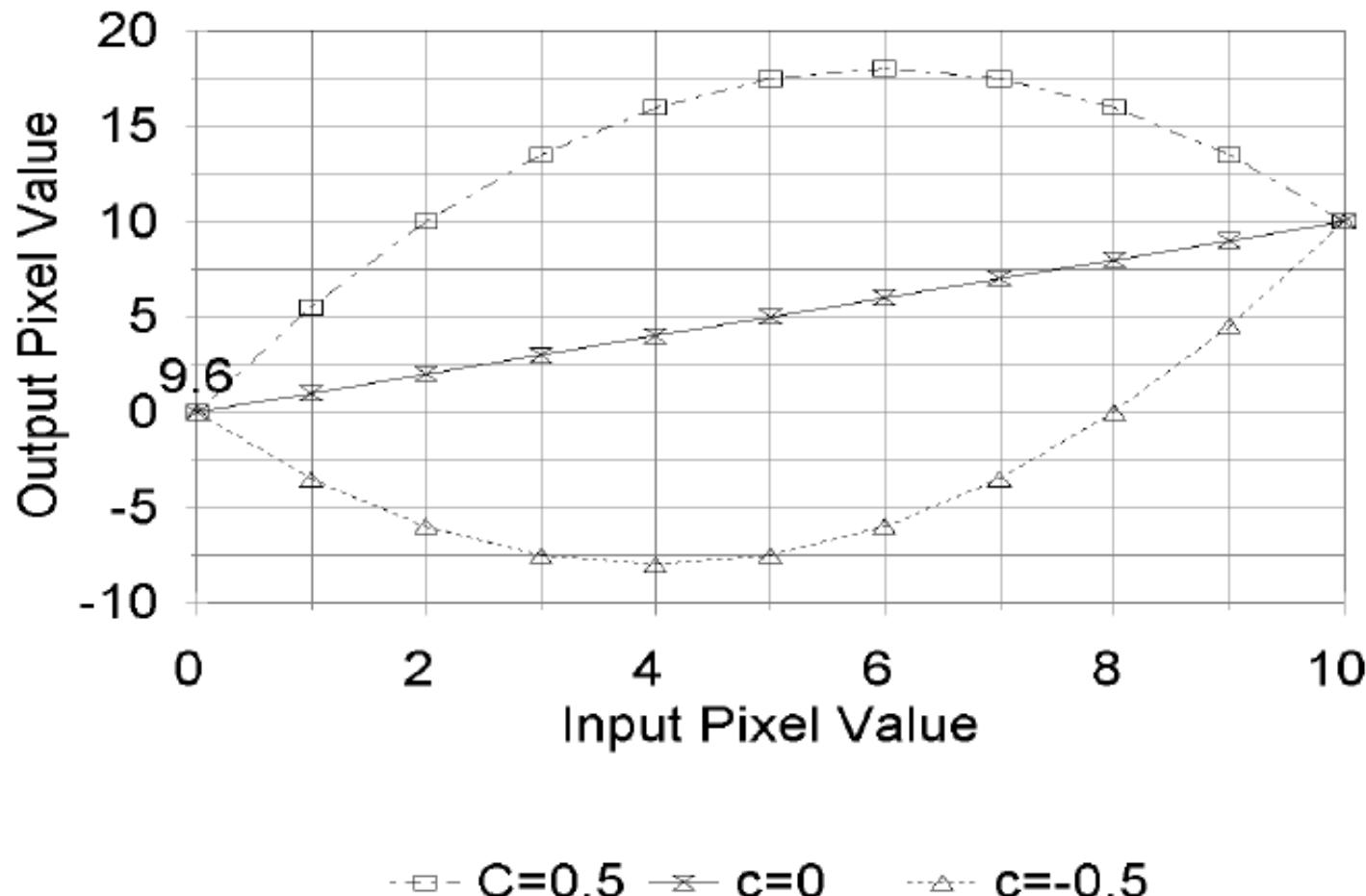
2.1 引言





第二章 常用的数学变换

2.1 引言





第二章 常用的数学变换

2.1 引言

2.2 空域变换

2.3 频率域变换

2.4 离散余弦变换

2.5 K-L 变换

2.6 其他正交变换



第二章 常用的数学变换

2.2 空域变换

2.2.1. 代数变换

图像的代数变换是指对两幅图像进行点对点的四则运算而得到一幅新的输出图像。图像的代数运算在图像处理中有着广泛的应用，它除了可以实现自身所需要的算术操作，还能为许多复杂的图像处理提供准备。

1) 加法运算

$$C(x, y) = A(x, y) + B(x, y)$$

2) 减法运算（差分）

$$C(x, y) = A(x, y) - B(x, y)$$



第二章 常用的数学变换

2.2 空域变换——2.2.1. 代数

□ 变换 加法运算

◆ 生成图像叠加效果

对于图像 $f(x, y)$ 和图像 $g(x, y)$ 求均值，有：

$$h(x, y) = \frac{1}{2}f(x, y) + \frac{1}{2}g(x, y)$$

会得到二次曝光的效果，其被推广为：

$$h(x, y) = \alpha \cdot f(x, y) + \beta \cdot g(x, y), \text{ 且 } \alpha + \beta = 1$$

我们可以得到各种图像合成的效果，也可以用两幅图像的衔接。

◆ 主要应用举例

- 去除“叠加性”噪声
- 生成图像叠加效果



第二章 常用的数学变换

2.2 空域变换——2.2.1. 代数

示例



+



=



加法运算





第二章 常用的数学变换

2.2 空域变换——2.2.1. 代数

□ 变换 减法运算

- ◆ 去除不需要的叠加性图案

设背景图像为 $b(x, y)$, 前景北京混合图像 $f(x, y)$, 则:

$$h(x, y) = f(x, y) - b(x, y)$$

其中, $h(x, y)$ 为去除了背景的图像

- ◆ 主要应用举例

- 去除不需要的叠加性图案
- 检测同一场景两幅图像之间的变化
- 计算物体边界的梯度



第二章 常用的数学变换

2.2 空域变换——2.2.1. 代数 变换



—



=



减法运算





第二章 常用的数学变换

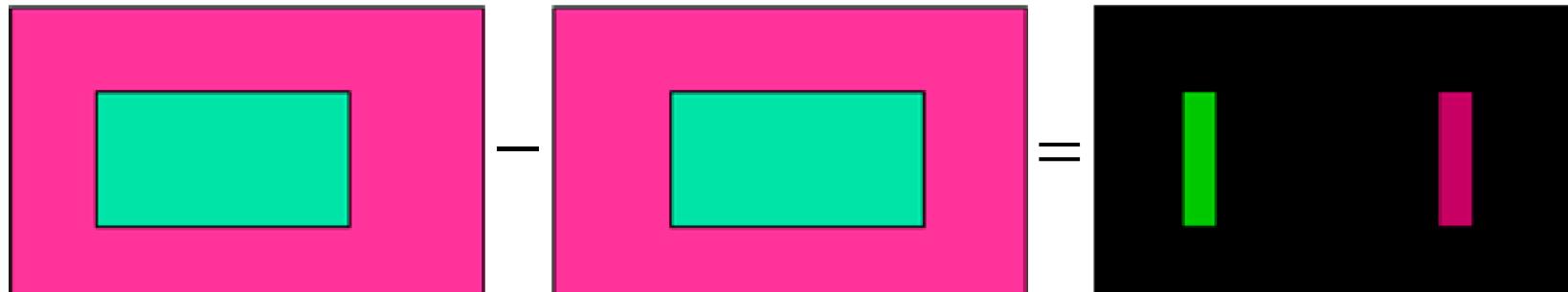
2.2 空域变换——2.2.1. 代数

□ 变换 减法运算

◆ 检测同一场景两幅图像之间的变化，设：

时间1的图像为 $t_1(x, y)$ ，时间2的图像为 $t_2(x, y)$

$$h(x, y) = t_1(x, y) - t_2(x, y)$$





第二章 常用的数学变换

2.2 空域变换 2.2.1. 代数变换

□ 乘法的定义

$$C(x, y) = A(x, y) \times B(x, y)$$

□ 主要应用举例

➤ 用二值模板图像的原图像做乘法





第二章 常用的数学变换

2.2 空域变换——2.2.1. 代数

□ 逻辑运算——求反

◆ 定义

$$h(x, y) = DN - f(x, y)$$

其中， DN 为 $f(x, y)$ 的灰度级

◆ 主要应用举例

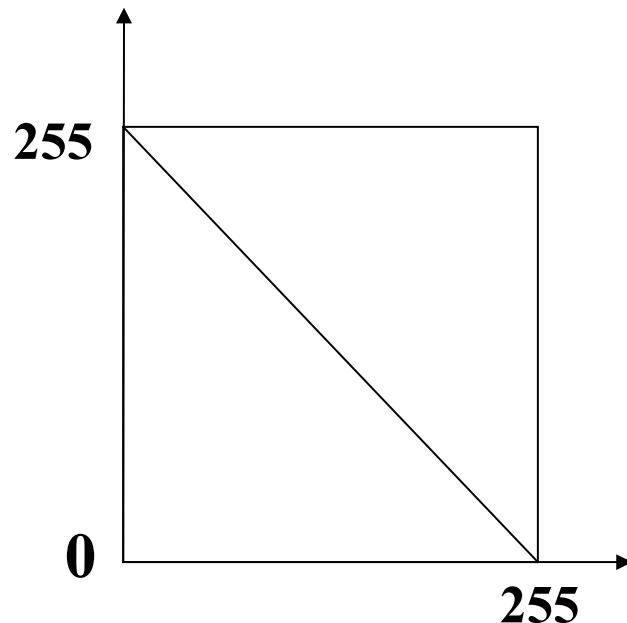
- 获得一个图像的负像
- 获得一个子图像的补图像
- 绘制区别于背景的、可恢复的图像



第二章 常用的数学变换

2.2 空域变换——2.2.1. 代数

变换
 逻辑运算——求反





第二章 常用的数学变换

2.2 空域变换——2.2.1. 代数

□ 逻辑运算——与

◆ 定义

$$h(x, y) = f(x, y) \wedge g(x, y)$$

◆ 主要应用举例

➤ 求两个子图像的相交子图



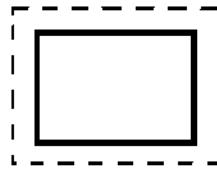


第二章 常用的数学变换

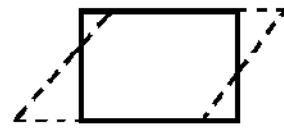
2.2 空域变换——2.2.2. 几何变换

图像在生成过程中，由于系统本身具有非线性或拍摄角度不同，会使生成的图像产生几何失真。几何失真一般分为**系统失真**和非**系统失真**。系统失真是有规律的、能预测的；非系统失真则是随机的。

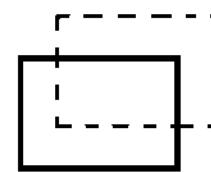
几何变换可以改变图像中物体之间的空间关系。这种运算可以看成是图像内的各物体在图像内移动的过程。例如，物体的转动、扭曲、倾斜、拉伸等等，都是几何运算的结果。如下图：



比例尺



歪斜



中心移动



扫描非线性



形状畸变



扭曲





第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何变换

□ 基本几何变换的定义

- ◆ 平移变换
- ◆ 旋转变换
- ◆ 镜像变换：水平镜像、垂直镜像
- ◆ 旋转变换
- ◆ 拉伸变换

□ 常用的基本几何变换

□ 数字图像几何变换的计算



第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何

变换

□ 基本几何变换的定义

◆ 设原图像为 $f(x, y)$

◆ 坐标变换函数设为：

$$x' = a(x, y) \quad y' = b(x, y)$$

◆ 因此，唯一确定了几何变换：

$$h(x', y') = f[a(x, y), b(x, y)]$$

其中， $h(x', y')$ 为目标图像



第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何变换

□ 基本几何变换——平移

已知原图像为 $f(x, y)$, 设平移量为 $(\Delta x, \Delta y)$, 则经平移后

的图像为 $h(x', y')$, 且有:

$$\begin{cases} x' = x + \Delta x \\ y' = y + \Delta y \end{cases}$$



第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何

变换

□ 基本几何变换——缩放

已知原图像为 $f(x, y)$, 设缩放系数分别为 a 和 b , 则经缩放后的图像为 $h(x', y')$, 且有:

$$\begin{cases} x' = ax \\ y' = by \end{cases}$$



第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何

变换

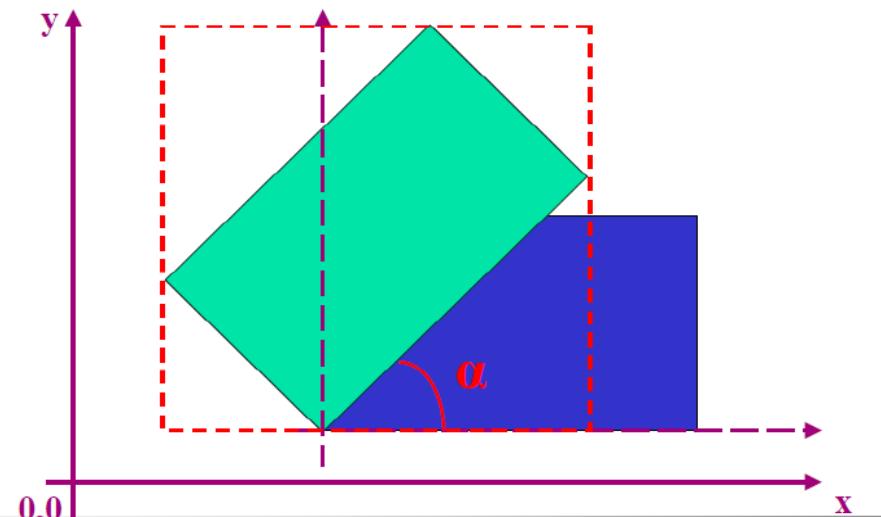
□ 基本几何变换——缩放

$$a(x, y) = x \cos(\theta) - y \sin(\theta)$$

$$b(x, y) = x \sin(\theta) + y \cos(\theta)$$

$$\begin{matrix} a(x, y) &= & \cos(\theta) & -\sin(\theta) & 0 \\ b(x, y) &= & \sin(\theta) & \cos(\theta) & 0 \\ 1 & & 0 & 0 & 1 \end{matrix}$$

■ 旋转变换: 绕原点旋转 α 度





第二章 常用的数学变换

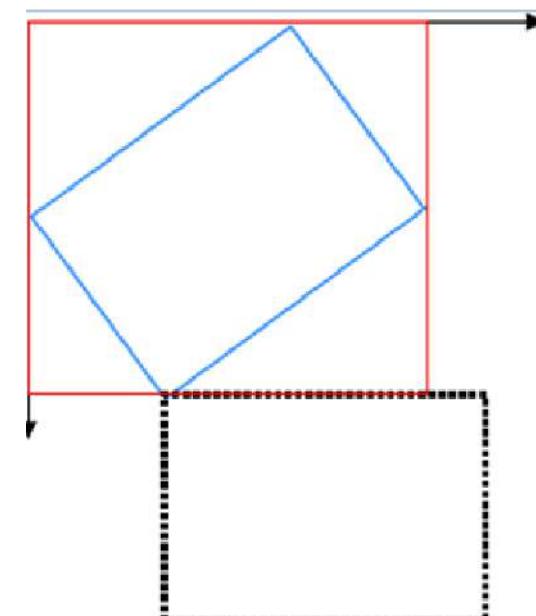
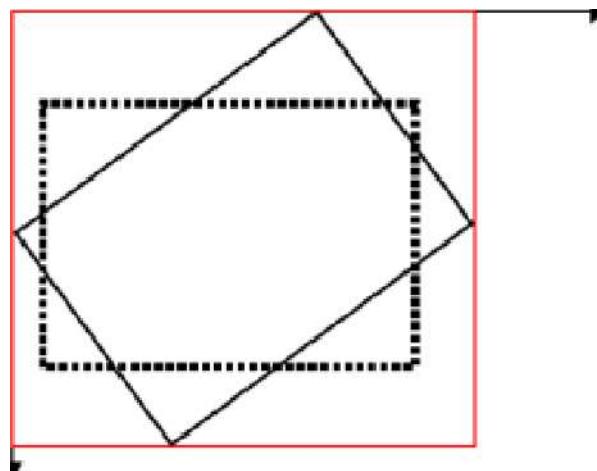
换

2.2 空域变换

2.2.2 几何变换

□ 基本几何变换——旋转变换的注意点 (1)

◆ 图像旋转之前，为了避免信息的丢失，一定有平移坐标，具体做法有如图所示的两种方法。





第二章 常用的数学变换

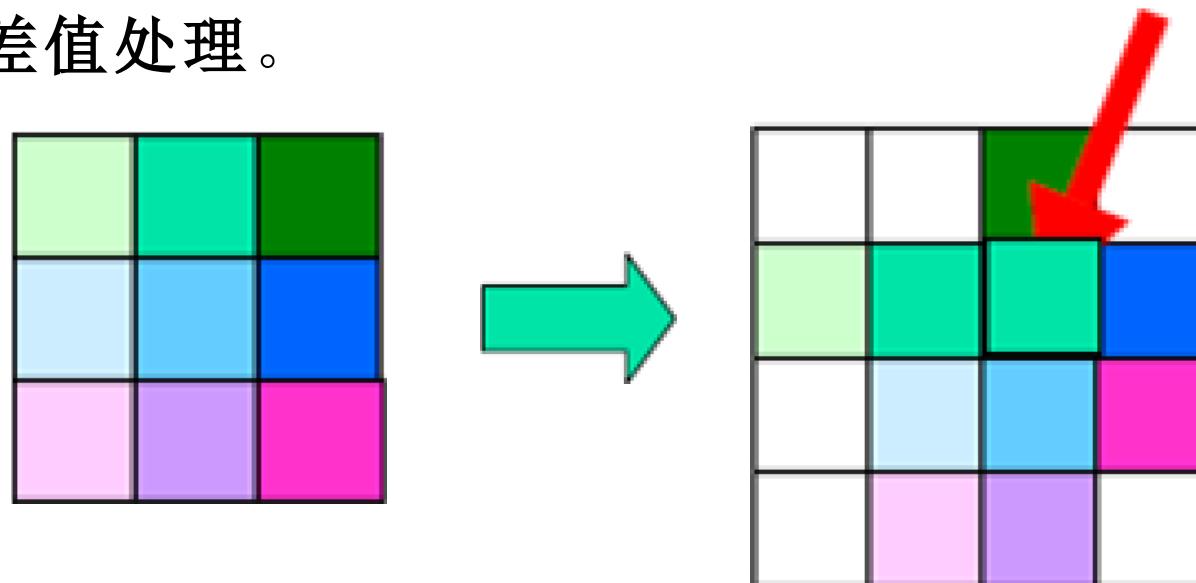
换

2.2 空域变换

2.2.2 几何变换

□ 基本几何变换——旋转变换的注意点 (2)

◆ 图像旋转之后，会出现许多的空洞点，对这些空洞点必须进行填充处理，否则画面效果不好，称这种处理为差值处理。





第二章 常用的数学变换

2.2 空域变换

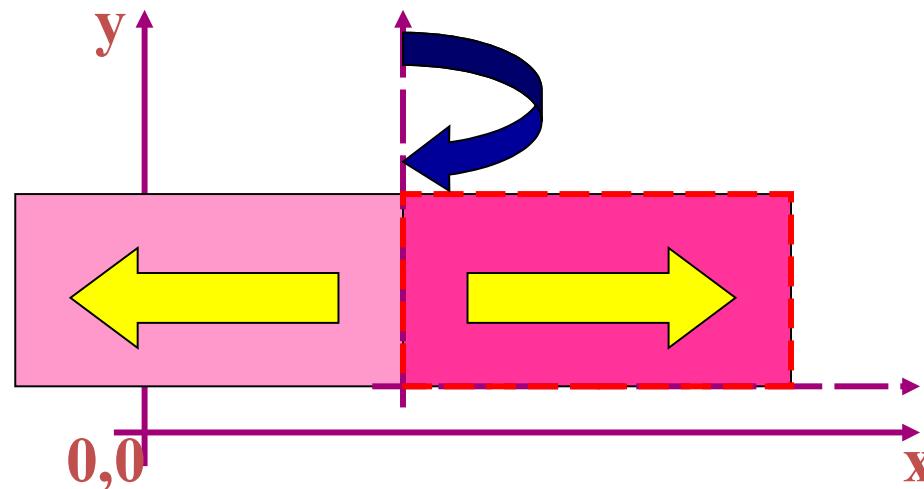
2.2.2 几何变换

□ 基本几何变换——水平镜像

已知原图像为 $f(x, y)$, 经水平镜像后的图像为 $h(x', y')$,

且有:

$$\begin{cases} x' = -x \\ y' = y \end{cases}$$





第二章 常用的数学变换

2.2 空域变换

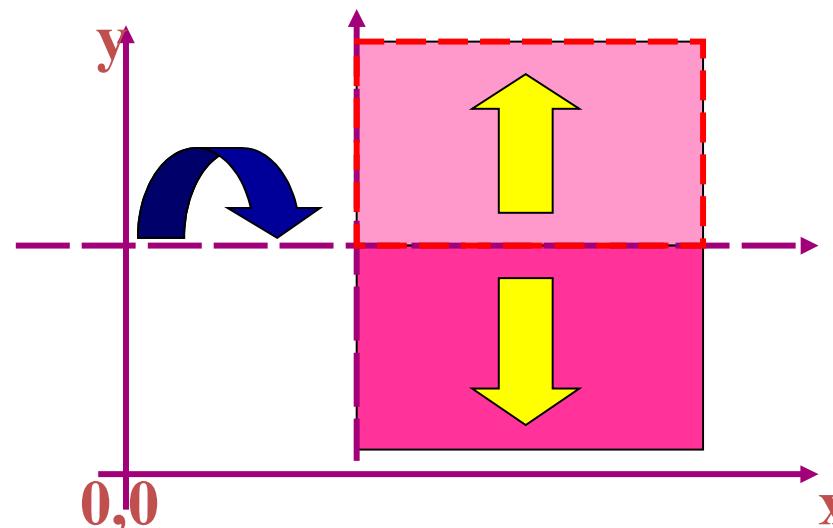
2.2.2 几何变换

□ 基本几何变换——垂直镜像

已知原图像为 $f(x, y)$, 经垂直镜像后的图像为 $h(x', y')$,

且有:

$$\begin{cases} x' = x \\ y' = -y \end{cases}$$





第二章 常用的数学变换

换

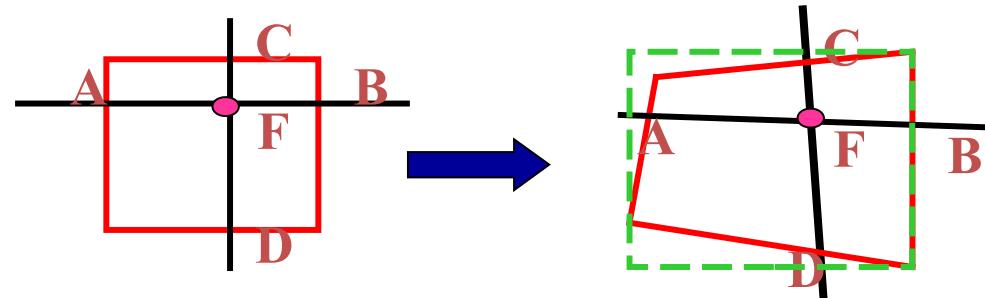
2.2 空域变换

2.2.2 几何变换

□ 基本几何变换——变形变换

下图显示了在失真和相应的校正图像中的四边形区域，四边的顶点是相应的“控制点”。假设四边形区域中的几何形变过程用双线性方程对来建模，即：

$$\begin{cases} x' = ax + by + cxy + d \\ y' = a'x + b'y + c'xy + d' \end{cases}$$





第二章 常用的数学变换

2.2 空域变换

2.2.2 几何变换

□ 基本激射变换的特征

◆ 坐标空间变换

- 范围发生变化
- 大小发生变化

◆ 像素值的变化

- 像素值不发生变化（位置变化）
- 像素值发生变化（旋转、缩放、变形变换）



第二章 常用的数学变换

2.2 空域变换

2.2.2 几何变换

□ 离散几何变换的计算问题

◆ 空间坐标

- 向前映射法
- 向后映射法

◆ 像素值计算——灰度插值（重采样）

- 最邻近插值法
- 双线性内插法（一阶插值）
- 高阶插值



第二章 常用的数学变换

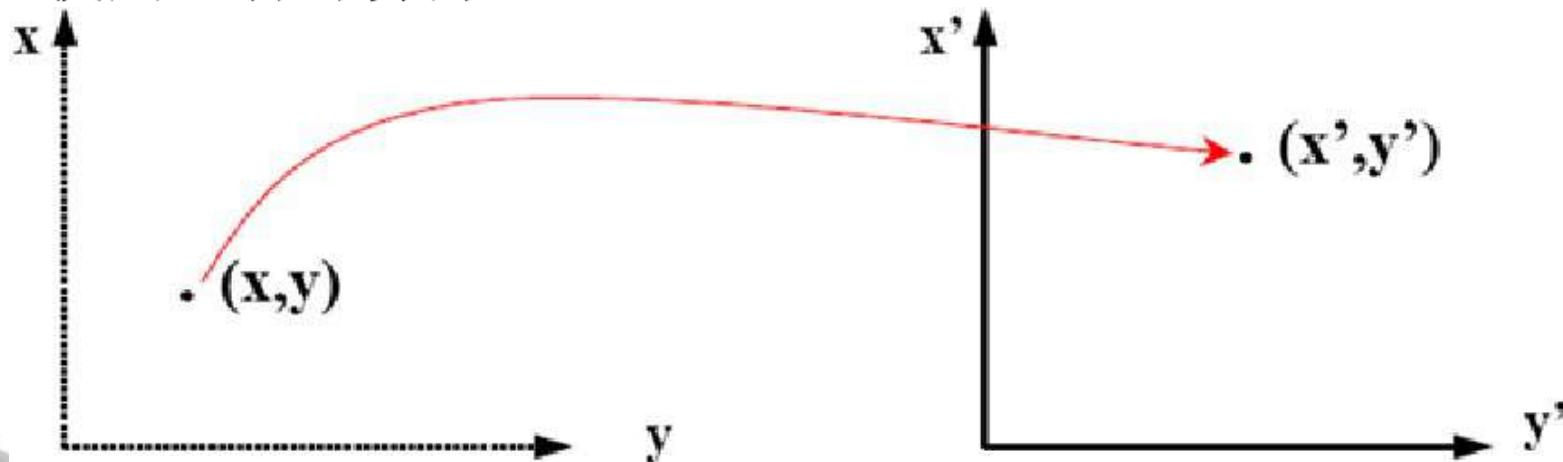
2.2 空域变换——2.2.2. 几何变换

□ 离散几何变换的计算问题

◆ 向前映射计算法

$$h(x', y') = f[a(x, y), b(x, y)]$$

从原图像坐标计算出目标图像坐标，镜像、平移变换使用这种计算方法





第二章 常用的数学变换

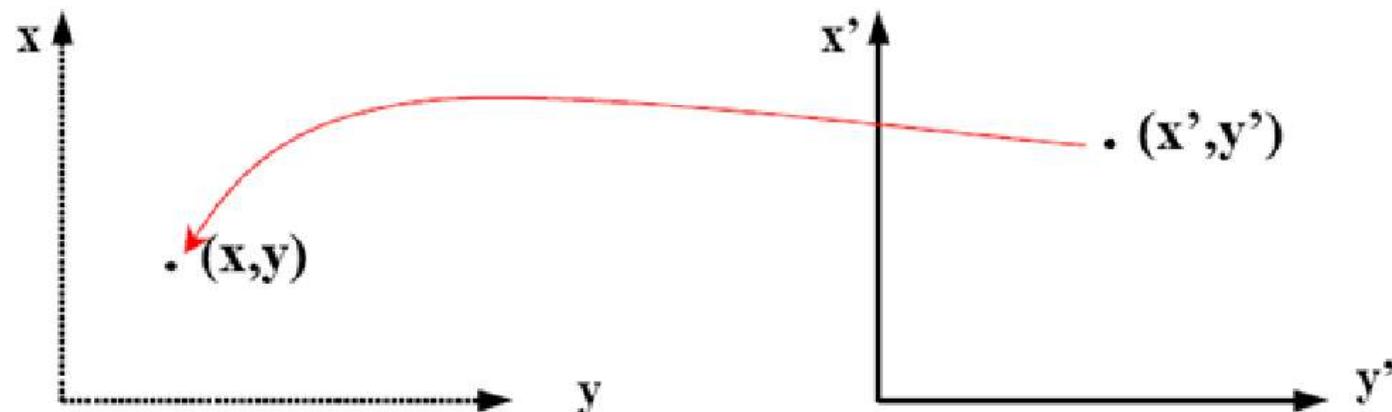
2.2 空域变换——2.2.2. 几何变换

□ 离散几何变换的计算问题

◆ 向后映射计算法

$$h[a'(x, y), b'(x, y)] = f(x, y)$$

从结果图像的坐标计算原图像的坐标，旋转、缩放、变形可以使用





国家级精品资源共享课

数字图像处理

(Digital Image Processing)

山东科技大学
江涛 教授



第二章 常用的数学变换

2.1 引言

2.2 空域变换

2.3 频率域变换

2.4 离散余弦变换

2.5 KL 变换

2.6 其他正交变换



第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何变换

□ 灰度插值——最近邻近插值法

◆ 选择最邻近点像素灰度值

(x', y') 点像素的灰度值为原图中 (x, y) 点的像素值

(x, y)		$(x+1, y)$	
	$\cdot (x', y')$		
$(x, y+1)$		$(x+1, y+1)$	



第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何

变换

□ 灰度插值——最近邻近插值法

- ◆ 简单快速
- ◆ 灰度保真性好
- ◆ 误差较大
- ◆ 视觉特性较差
- ◆ 马赛克效应





第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何变换

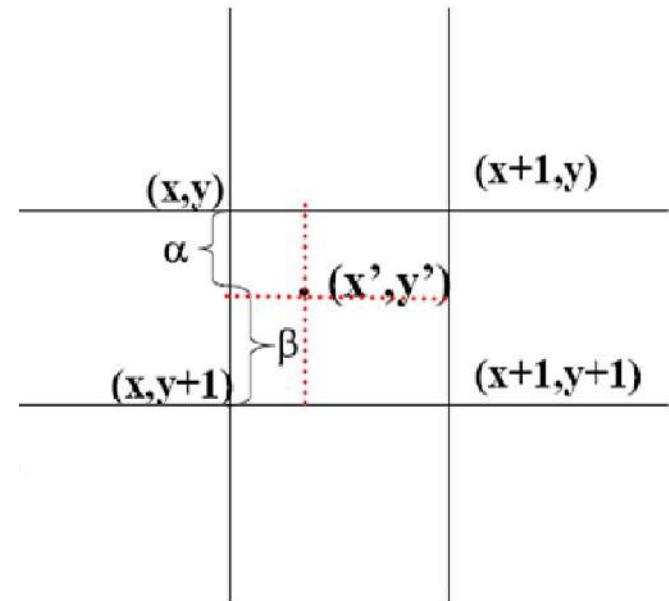
□ 灰度插值——双线性插值（一阶插值）

$$f'(x',y') = a \cdot f(x,y) + b \cdot f(x,y+1)$$

$$f''(x',y') = c \cdot f(x,y) + d \cdot f(x+1,y)$$

$$f'''(x',y') = u \cdot f(x,y+1) + v \cdot f(x+1,y+1)$$

$$f^{(4)}(x',y') = w \cdot f(x+1,y) + z \cdot f(x+1,y+1)$$





第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何变换

□ 双线性插值——简化计算方法

◆ 应用双曲抛物面方程

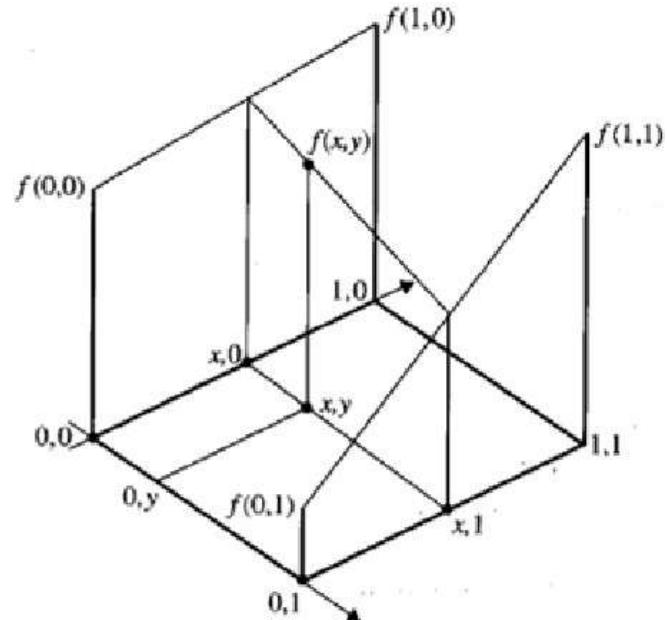
$$f(x, y) = ax + by + cxy + d$$

◆ 归一化坐标值

$$0 < x < 1, 0 < y < 1$$

◆ 可有

$$\begin{aligned} f(x, y) = & [f(1,0) - f(0,0)]x + [f(0,1) - f(0,0)]y \\ & + [f(1,1) + f(0,0) - f(0,1) - f(1,0)]xy \\ & + f(0,0) \end{aligned}$$





第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何

变换

□ 灰度插值——双线性插值(一阶插值)

◆ 双线性插值一般理论——双曲坡屋面方程差值

$$f(x, y) = ax + by + cxy + d$$

◆ 需得到4个位置参数——利用4个已知点

◆ 特点

➤ 计算中较为充分地考虑相邻各点的特征，具有灰度平滑过度特点

➤ 一般情况下可得到满意结果

➤ 具有低通滤波特性，使图像轮廓模糊

➤ 平滑作用使图像细节退化，尤其在放大时

➤ 不连续性会产生不希望的结果

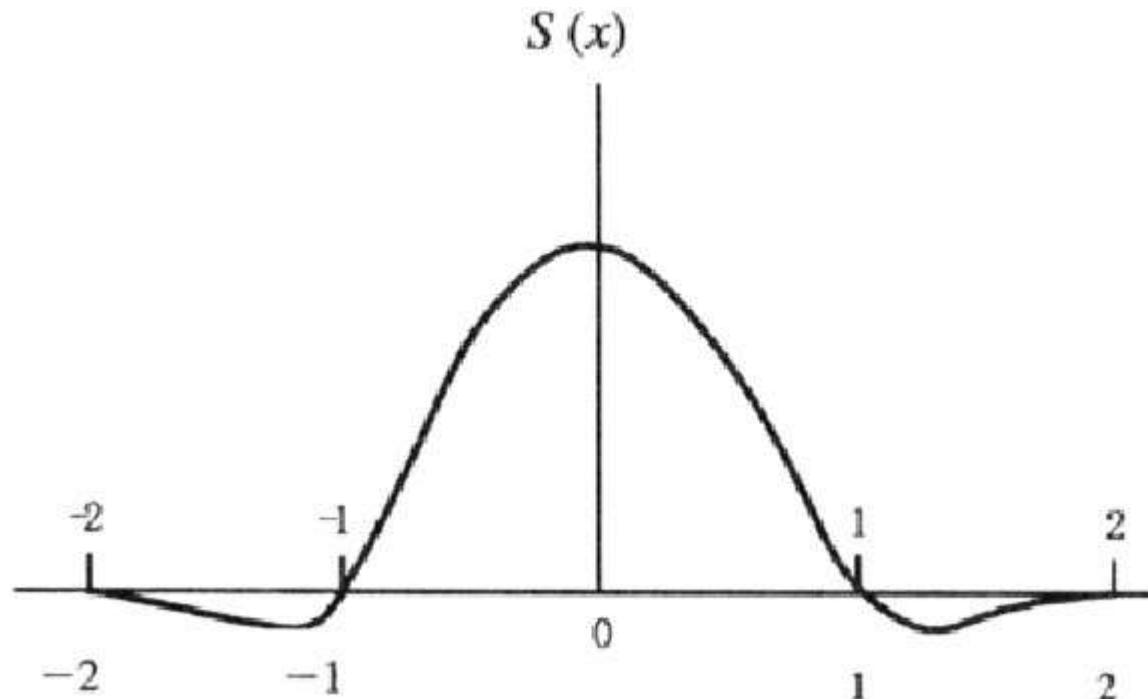


第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何变换

□ 灰度插值——高阶插值

◆ 如果简化计算，仅取原点周围有限范围函数：





第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何变换

□ 灰度插值——高阶插值

◆ 利用三次多项式来近似理论上的最佳插值函数 $\text{sinc}(x)$:

$$S(x) = \begin{cases} 1 - 2|x|^2 + |x|^3 & |x| < 1 \\ 4 - 8|x| + 5|x|^2 - |x|^3 & 1 \leq |x| \leq 2 \\ 0 & |x| > 2 \end{cases}$$

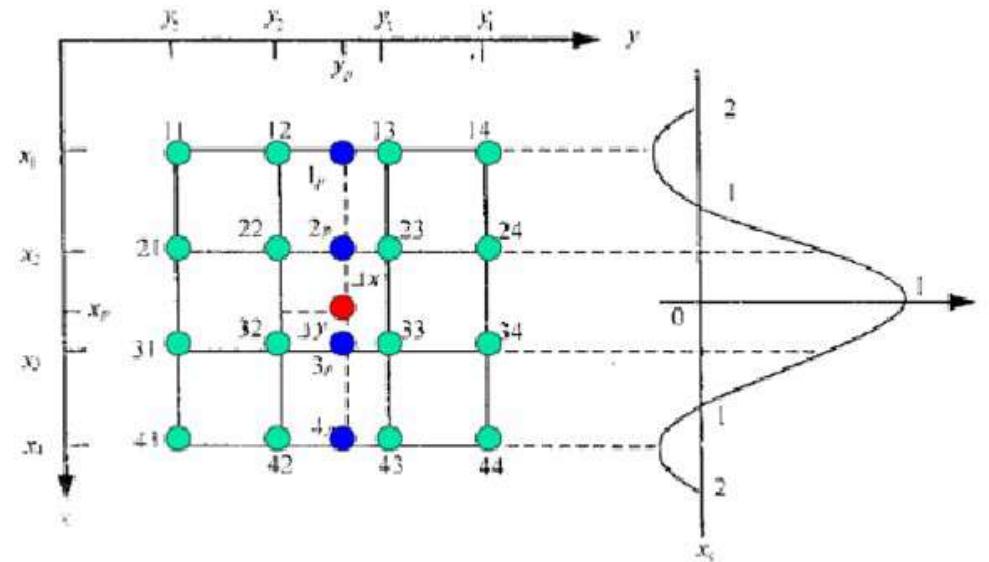
◆ 由此形成常用的三次卷积插值算法，又称三次内插法



第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何变换

- 灰度插值——三次卷积插值算法实现
 - ◆ 利用待插值点周围的16个邻点像素值：



- ◆ 确定辅助点位 $1p, 2p, 3p, 4p$ 各点亮度值，从而确定 p 点亮度值



第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何变换

□ 灰度插值——三次卷积插值算法的特点

- ◆ 为满足二维Nyquist条件下，最佳重构公式的近似；
- ◆ 只有图像满足特点的条件，三次卷积插值算法才可获得最佳结果；
- ◆ 可使待求点的灰度值更好地模拟实际可能值；
- ◆ 可取得更好的视觉效果；
- ◆ 三次卷积内插突出的优点是高频信息损失少，可将噪声平滑；
- ◆ 4×4 时，像元均值和标准差信息损失小；
- ◆ 计算量大为增加；



第二章 常用的数学变换

2.2 空域变换——2.2.2. 几何变换

换

□ 灰度插值——图像处理中内插方法的选择

内插方法的选择除了考虑图像的现实要求及计算量，还要考虑内插结果对分析的影响

- ◆ 当纹理信息为主要信息时，最邻近采样将严重改变原图像的纹理信息
- ◆ 但灰度信息为主要信息时，双线性内插及三次卷积插值将减少图像异质性，增加图像同质性，其中，双线性内插方法使这种变化更明显



第二章 常用的数学变换

2.1 引言

2.2 空域变换

2.3 频率域变换

2.4 离散余弦变换

2.5 KL 变换

2.6 其他正交变换



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）



傅立叶 (Fourier, Jean Baptiste Joseph, 1768-1830)

- 法国数学家及物理学家
- 最早使用定积分符号，改进符号法则及根数判别方法。

● 傅立叶级数（三角级数）创始人

1822 年在代表作《热的分析理论》中解决了热在非均匀加热的固体中分布传播问题，成为分析学在物理中应用的最早例证之一，对 19 世纪数学和理论物理学的发展产生深远影响。傅立叶级数（即三角级数）、傅立叶分析等理论均由此创始。



第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.1 总论傅立叶变换的由来

傅立叶级数



傅立叶积分



傅立叶变换

1807年向巴黎科学院呈交《热的传播》论文，推导出著名的热传导方程，并在求解该方程时发现解函数可以由三角函数构成的级数形式表示，从而提出任一函数都可以展成三角函数的无穷级数。

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx dx \quad (n = 0, 1, 2, \dots)$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin nx dx \quad (n = 0, 1, 2, 3)$$



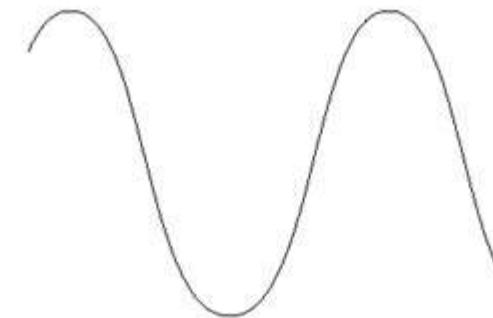
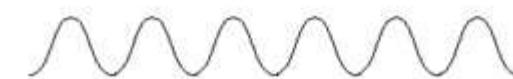
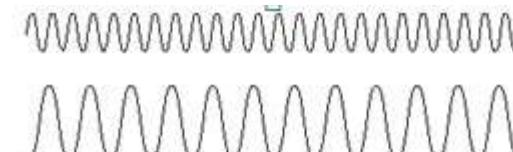
第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.1 总论

□ 傅立叶级数的物理含义

任何时域连续的复杂波形，通过傅里叶变换，都可以变成一系列简谐波之和

。





第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.2 基本概念 □ 一维傅立叶变换

函数的一维连续傅立叶变换式定义为 $F(u) = \int_{-\infty}^{\infty} f(t) e^{-j2\pi ut} dt$

其反变换为: $f(t) = \int_{-\infty}^{\infty} F(u) e^{j2\pi ut} du$

这里的函数必须满足只有有限个间断点、有限个极值和绝对可积的条件，并且也应是可积的。 F 一般是实函数，而是复函数，它由实部和虚部组成:

$$F(u) = R(u) - jI(u) = |F(u)| e^{j\varphi(u)}$$

$$|F(u)| = \sqrt{R^2(u) + I^2(u)}, \quad \varphi(u) = \arctan \frac{I(u)}{R(u)}$$

其中, $|F(u)|$ 称为频谱,
称为相位角。



第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.2 基本概念 □ 一维傅立叶变换

◆ 一维离散傅立叶变换定义为:

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-j2\pi ux/N} \quad u = 0, 1, \dots, N-1$$

◆ 其反变换为:

$$f(x) = \sum_{u=0}^{N-1} F(u) e^{j2\pi ux/N} \quad x = 0, 1, \dots, N-1$$



第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.2 基本概念 三维傅立叶变换

◆假设以矩形网格采样得到的图像用 来表示
， 则 的二维离散傅立叶变换可以表示为：

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})} \quad u = 0, 1, \dots, M-1; v = 0, 1, \dots, N-1$$

◆其反变换为：

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(\frac{ux}{M} + \frac{vy}{N})} \quad x = 0, 1, \dots, M-1; y = 0, 1, \dots, N-1;$$



第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.2 基本概念 □ 二维傅立叶变换

与一维离散傅立叶变换的情况类似，可以定义二维离散傅立叶变换的频谱和相位角如下：

$$F(u, v) = |F(u, v)| e^{j\varphi(u, v)} = R(u, v) + jI(u, v)$$

$$|F(u, v)| = \sqrt{R^2(u, v) + I^2(u, v)}, \quad \varphi(u, v) = \arctan \frac{I(u, v)}{R(u, v)}$$



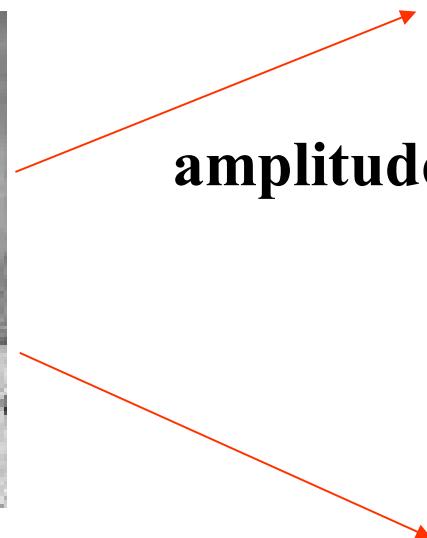
第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.3 显示

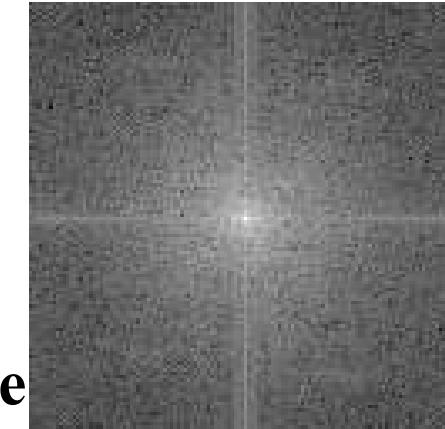


original

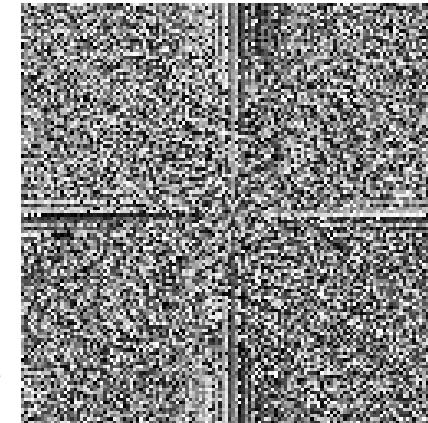
功率谱和相位谱



amplitude



phase

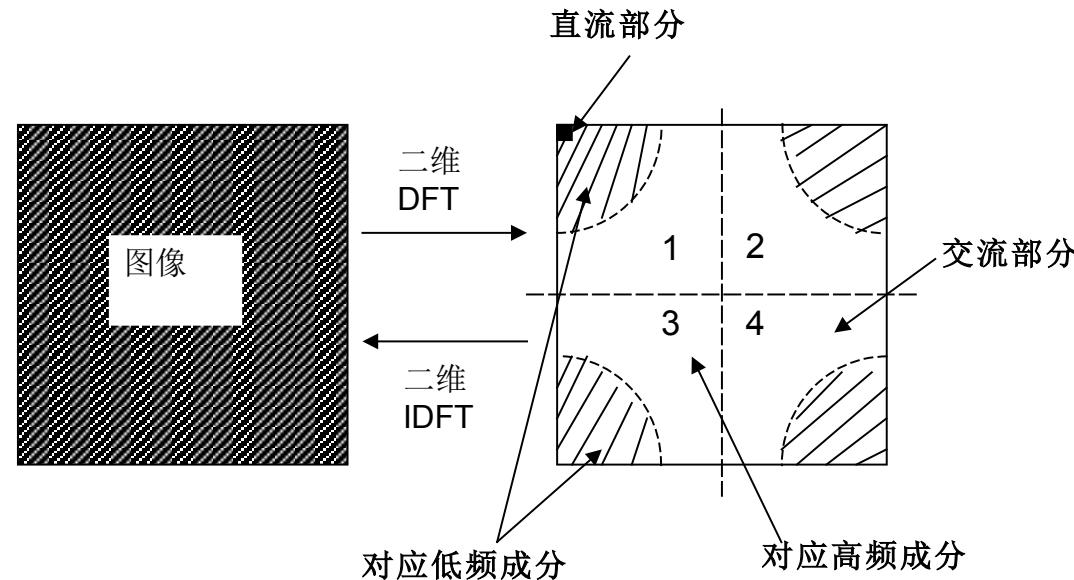




第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.3

显示图像的二维离散傅立叶变换的频率成分分布





国家级精品资源共享课

数字图像处理

(Digital Image Processing)

山东科技大学
江涛 教授



第二章 常用的数学变换

2.1 引言

2.2 空域变换

2.3 频率域变换

2.4 离散余弦变换

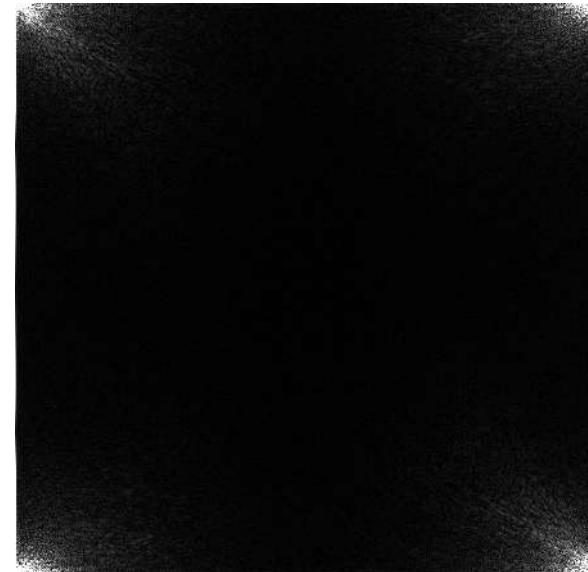
2.5 KL 变换

2.6 其他正交变换



第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.3 显示



图像及其频谱图像示意图



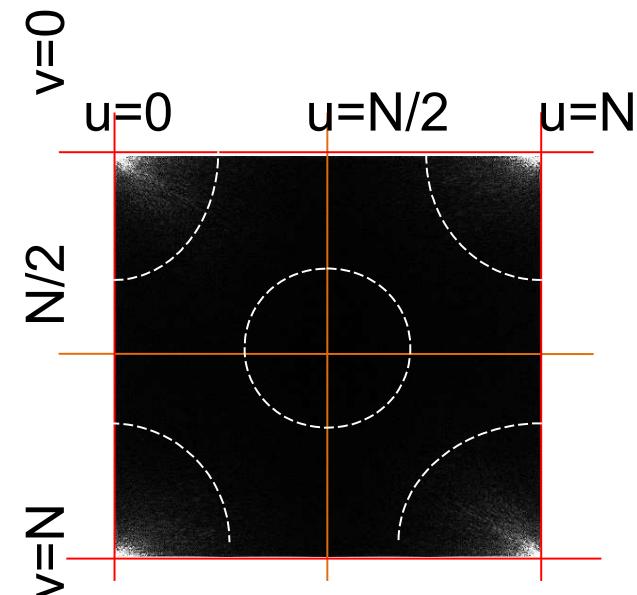
第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.3 显示

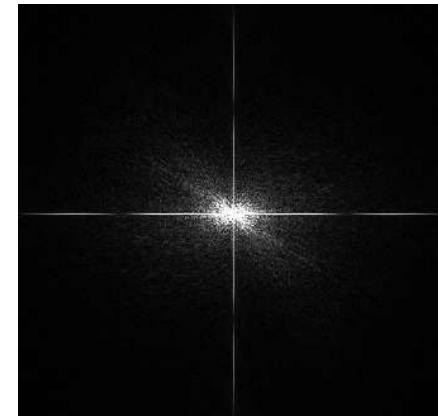
零点漂移



原始图像



零点
漂移
图像
的功
率谱



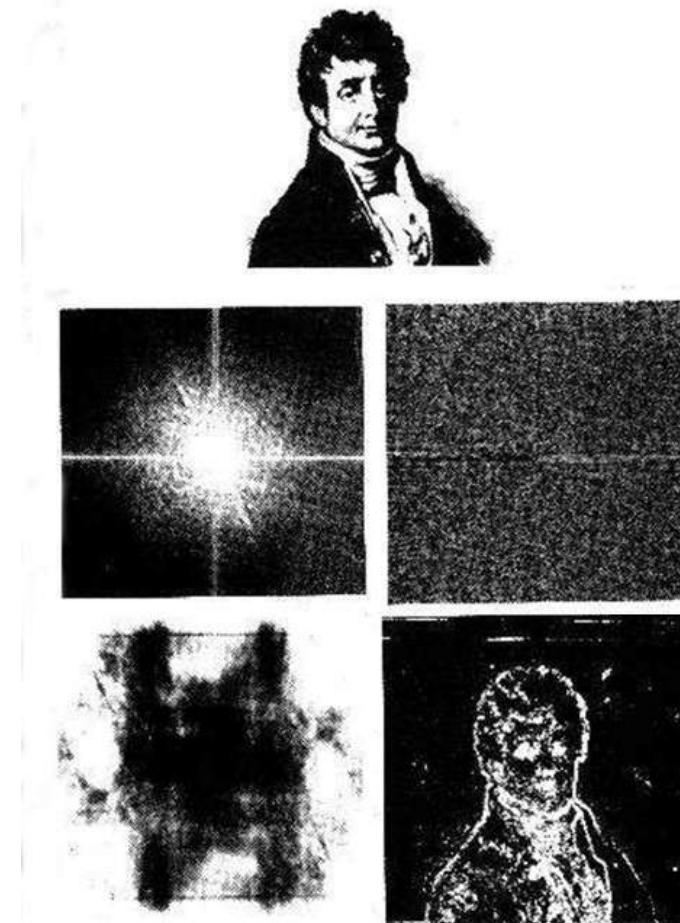


第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.3 显示

- 输入图像
- 幅值谱
- 相位谱
- 由幅值谱重构的图象
- 由相位谱重构的图象

结论：相位谱可能具有更重要的应用

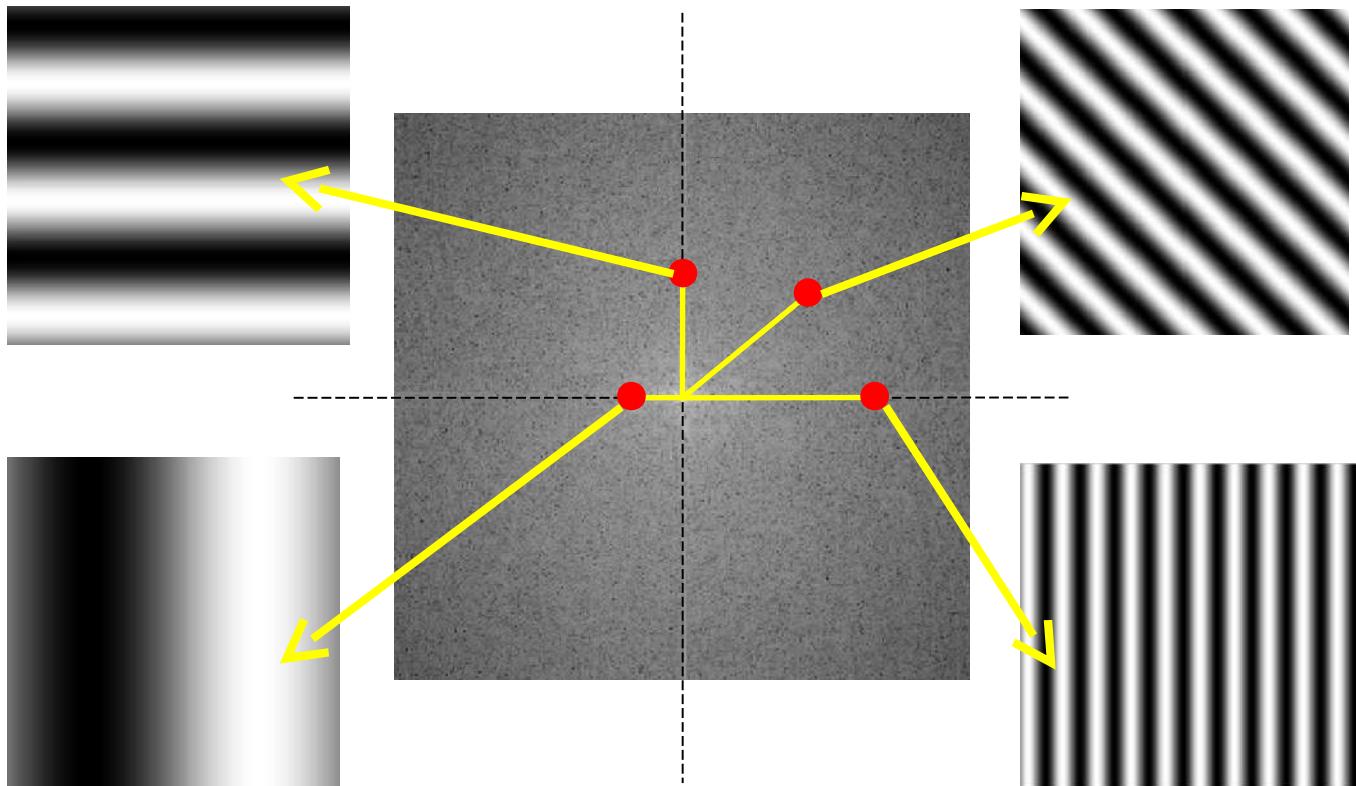




第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.3 显示

图像域功率谱关系

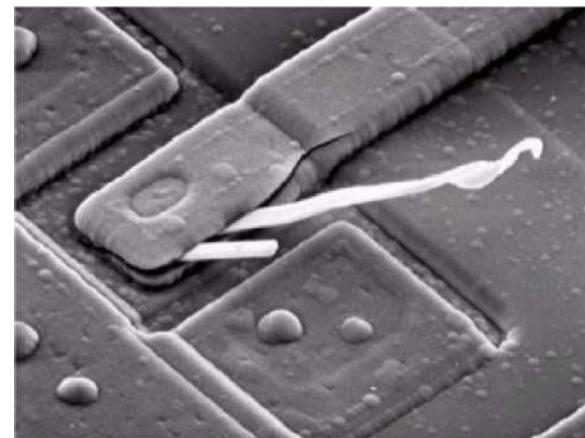




第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.3 显示

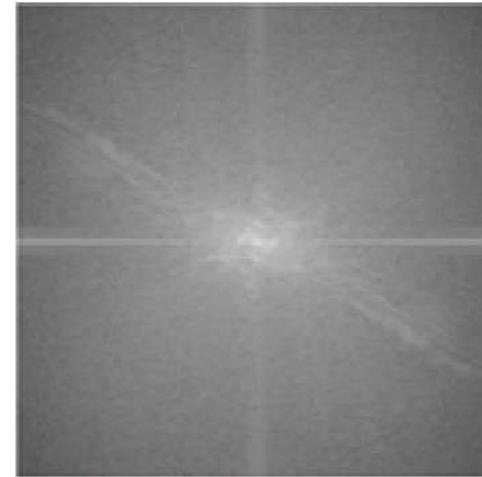
图像傅立叶变换示例：注意
低频（背景）和高频（细节、
边缘）
信息的分布



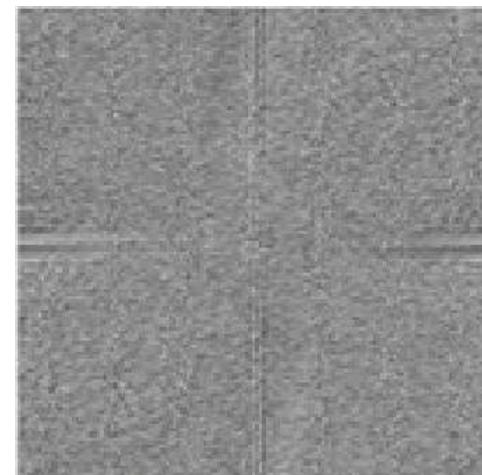


第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.3 显示



Magnitude spectrum



Phase spectrum





第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.3 显示



Magnitude spectrum



Phase spectrum





第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.4 基本性质

□ 可分离性

◆ 正变换:

$$F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} e^{-j2\pi x/N} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi y/N} = F_y\{F_x[f(x, y)]\}$$

◆ 反变换:

$$\begin{aligned} f(x, y) &= \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(ux+vy)/N} = \sum_{u=0}^{N-1} e^{j2\pi ux/N} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi vy/N} \\ &= F_u^{-1}\{F_v^{-1}[F(u, v)]\} \end{aligned}$$



第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.4 基本性质

□ 可分离性

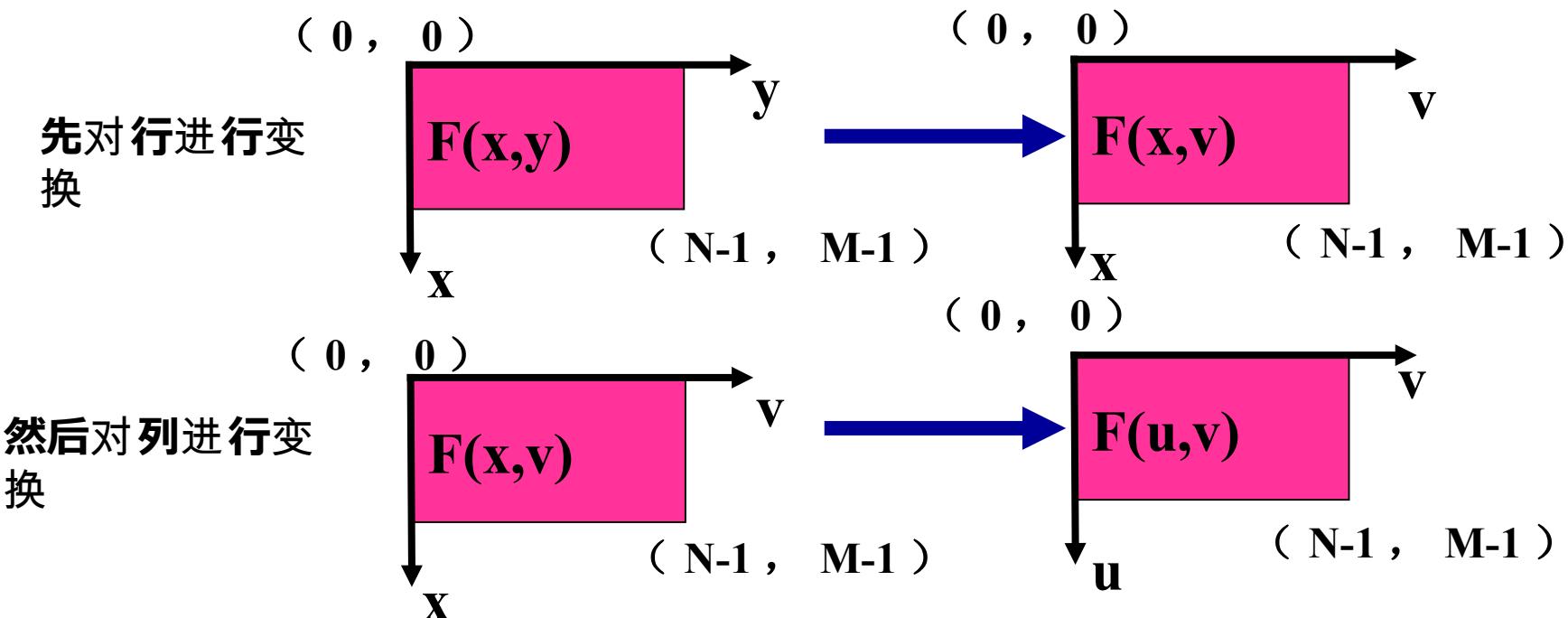
由可分离性可知，一个二维傅立叶变换可分解为两步进行，其中每一步都是一个一维傅立叶变换。先对 $f(x, y)$ 按行进行傅立叶变换得到 $F(x, v)$ ，再对 $F(x, v)$ 按列进行傅立叶变换，便可得到 $f(x, y)$ 的傅立叶变换结果 $F(u, v)$ 。显然对 $f(x, y)$ 先按列进行离散傅立叶变换，再按行进行离散傅立叶变换也是可行的。



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.4 基本性质

□ 可分离性





第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.4 基本性质

□ 可分离性

由可分离性可知，一个二维傅立叶变换可分解为两步进行，其中每一步都是一个一维傅立叶变换。先对 $f(x, y)$ 按行进行傅立叶变换得到 $F(x, v)$ ，再对 $F(x, v)$ 按列进行傅立叶变换，便可得到 $f(x, y)$ 的傅立叶变换结果 $F(u, v)$ 。显然对 $f(x, y)$ 先按列进行离散傅立叶变换，再按行进行离散傅立叶变换也是可行的。



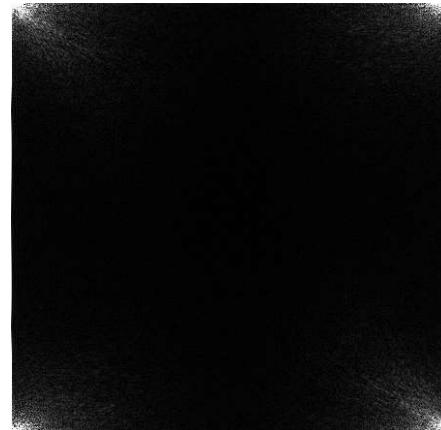
第二章 常用的数学变换

2.3 频率域变换(傅立叶变换)——2.3.4 基本性质

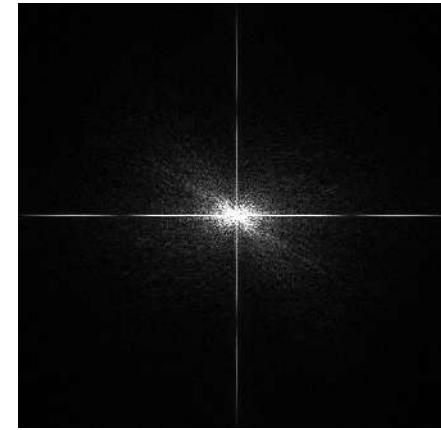
□ 平移性



(a) Lena 原图像
傅立叶频谱



(b) 无平移的傅立叶频谱



(c) 平移后的



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.4 基本性质

□ 线性特

征

$$f_1(x, y) + f_2(x, y) \Leftrightarrow F_1(u, v) + F_2(u, v)$$



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.4 基本性质

□周期性

傅立叶变换和反变换均以为 N 周期，即：

$$F(u, v) = F(u + N, v) = F(u, v + N) = F(u + N, v + N)$$

傅立叶变换的周期性表明，尽管 $F(u, v)$ 对无穷多个 u 和 v 的值重复出现，但只需根据在任意周期内的 N 个值就可以从 $F(u, v)$ 得到 $f(x, y)$ 。也就是说，只需一个周期内的变换就可以将 $F(u, v)$ 完全确定。这一性质对于 $f(x, y)$ 在空域里也同样成立。





第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.4 基本性质

□ 共轭对称

如果 $f(x, y)$ 是实函数，则它的傅立叶变换具有共轭对称性：

$$F(u, v) = F^*(-u, -v)$$

$$|F(u, v)| = |F(-u, -v)|$$

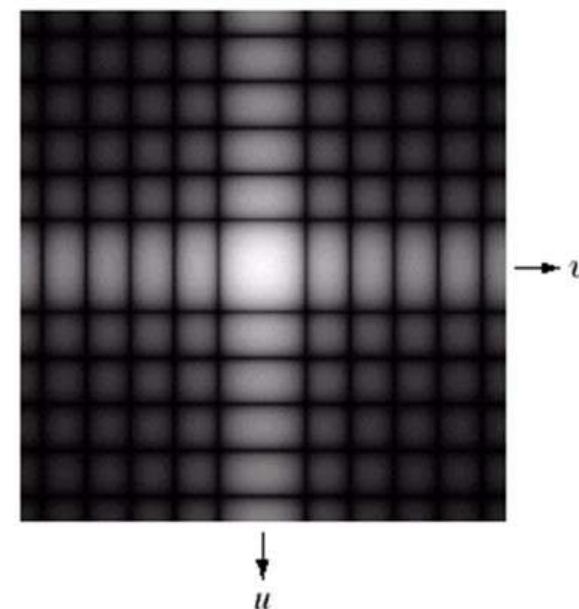
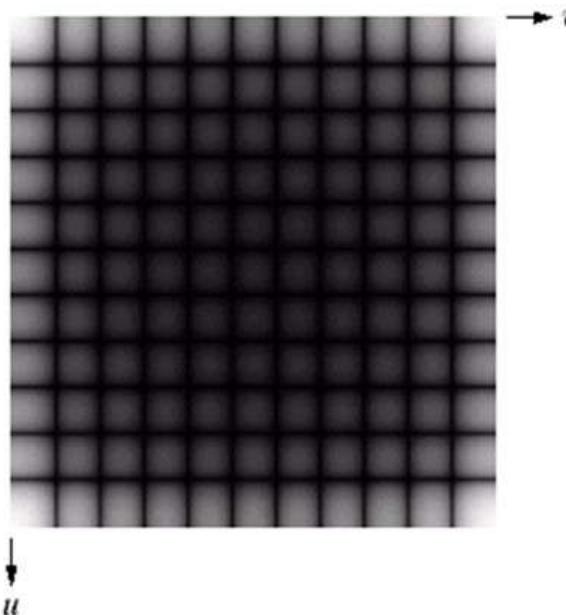
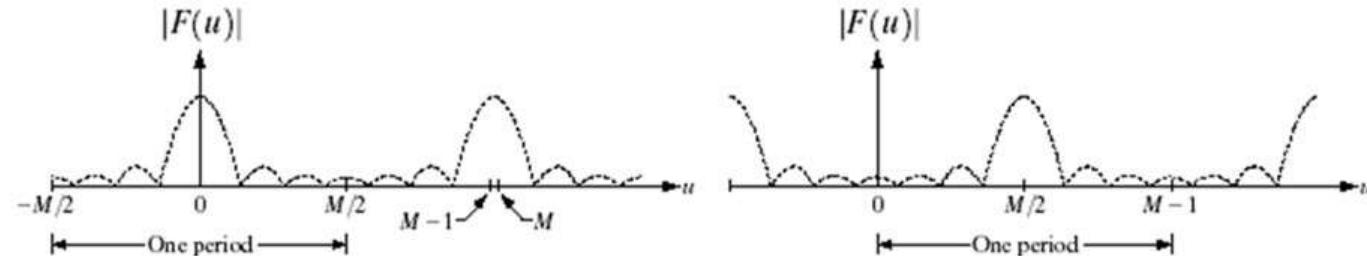
其中， $F^*(u, v)$ 是 $F(u, v)$ 的复共轭。

图像傅立叶变化的周期性和共轭对称性对图像的频谱分析和显示带来很大的益处。



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.4 基本性质





第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.4 基本性质

□ 旋转不变

僅引入极坐标使
$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases} \quad \begin{cases} u = \omega\cos\varphi \\ v = \omega\sin\varphi \end{cases}$$

则 $f(x, y)$ 和 $F(u, v)$ 分别表示为 $f(r, \theta)$ 和 $F(\omega, \varphi)$ 。

$$f(r, \theta + \theta_0) \Leftrightarrow F(\omega, \varphi + \theta_0)$$

如果图像本身在空间域上旋转，则其二维离散傅立叶变换在频率域上也会旋转，而且旋转的方向和角度相同。

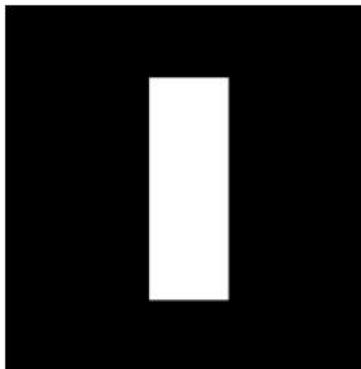


第二章 常用的数学变换

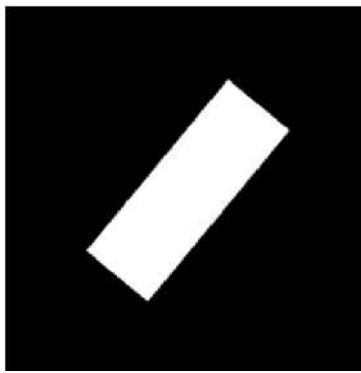
2.3 频率域变换（傅立叶变换）——2.3.4 基本性质

□ 旋转不变性

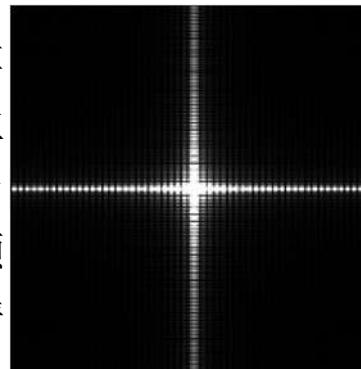
原图像



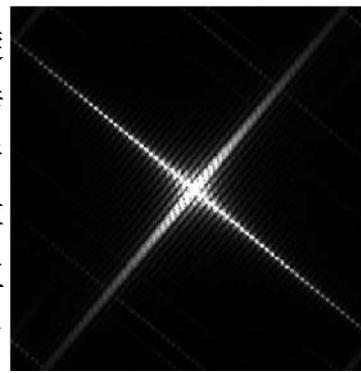
旋转后图像



傅立叶频谱



旋转后傅立叶





第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.4 基本性质

□ 比例性质

$$af(x, y) \Leftrightarrow aF(u, v)$$

$$f(ax, by) \Leftrightarrow \frac{1}{|ab|} F\left(\frac{u}{a}, \frac{v}{b}\right)$$

在空间域比例尺度的展宽，对应于在频域比例尺度的压缩，其幅值也减少为原来的 $\frac{1}{|ab|}$ 。当 $a = b = -1$ 时：

$$f(-x, -y) \Leftrightarrow F(-u, -v)$$

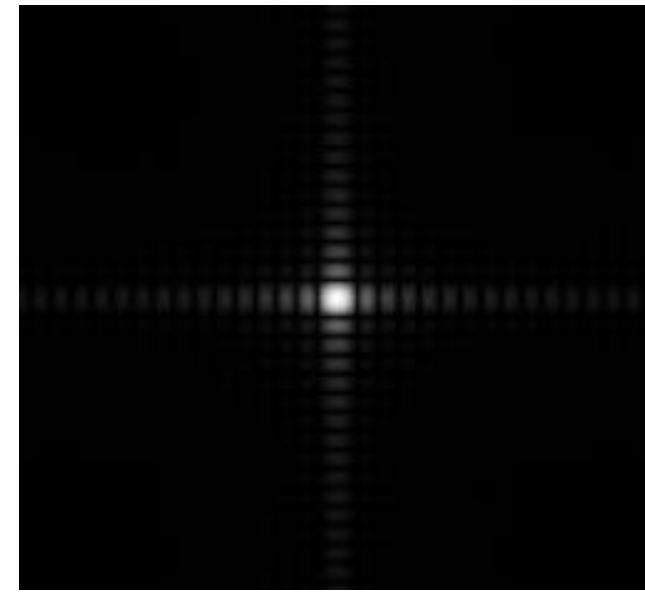
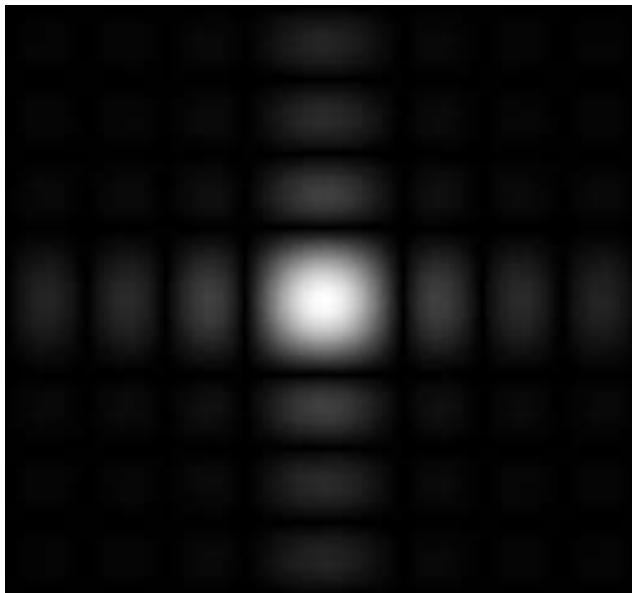


第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.4 基本性质

□ 比例性

质



a) 比例尺度展宽前的频谱

b) 比例尺度展宽后的频谱



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.4 基本性质

□ 平均值性质

◆ 二维离散函数 $f(x, y)$ 的平均值定义为：

$$\bar{f}(x, y) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)$$

◆ 二维离散函数 $f(x, y)$ 的傅立叶变换的平均值定义为：

$$F(0, 0) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)$$

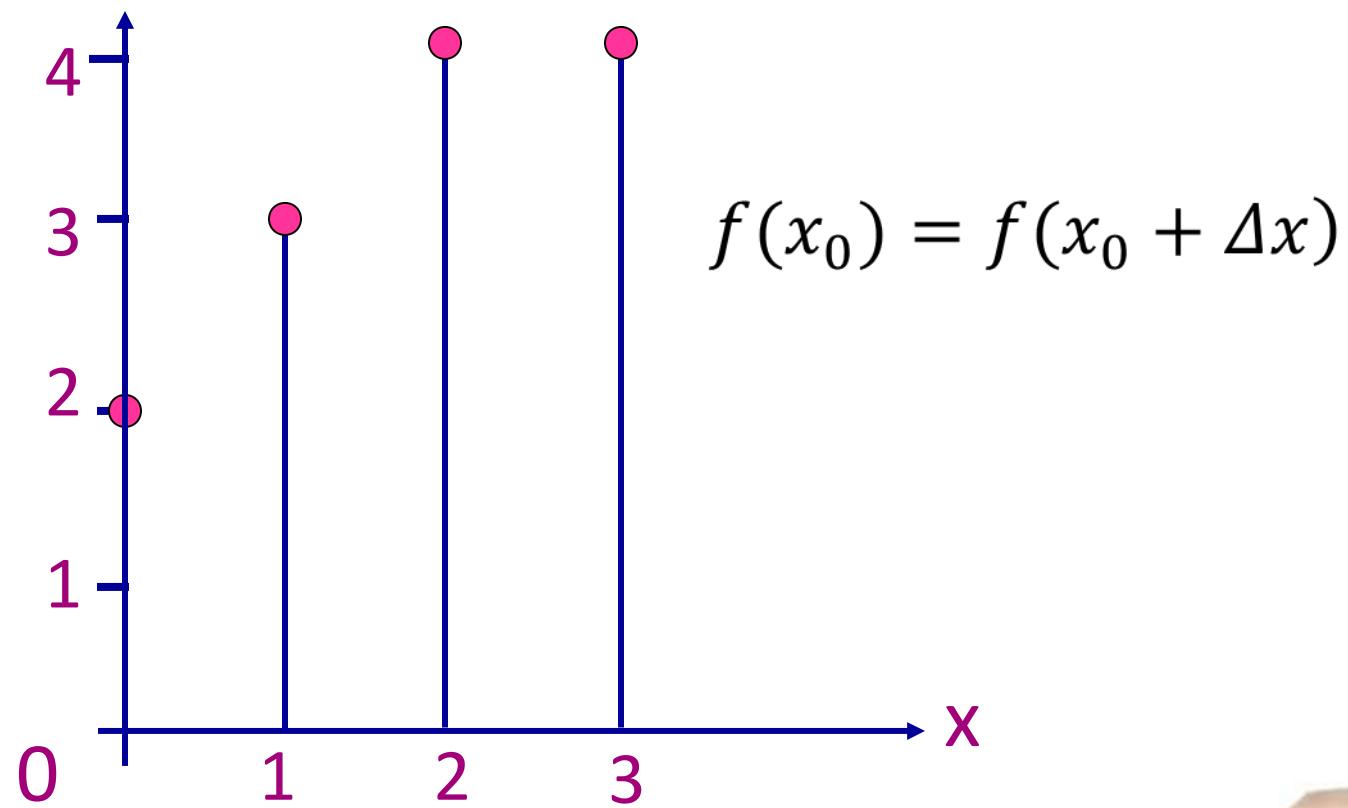
则有：

$$\bar{f}(x, y) = F(0, 0)$$



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.5 计算实例





第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.5 计算实例

$$F(0) = \frac{1}{4} \sum f(x) \exp[0]$$

$$= \frac{1}{4} [f(0) + f(1) + f(2) + f(3)]$$

$$= \frac{1}{4} (2 + 3 + 4 + 4)$$

$$= 3.25$$

$$F(1) = \frac{1}{4} \sum f(x) \exp[-j2\pi x/4]$$

$$= \frac{1}{4} \left(2e^0 + 3e^{-j2\pi^2/4} + 4e^{-j2\pi^3/4} \right)$$

$$= \frac{1}{4} (-2 + j)$$

$$F(2) = -1/4(1 + j^0)$$

$$F(3) = -1/4(2 + j)$$



国家级精品资源共享课

数字图像处理

(Digital Image Processing)

山东科技大学
江涛 教授



第二章 常用的数学变换

2.1 引言

2.2 空域变换

2.3 频率域变换

2.4 离散余弦变换

2.5 KL 变换

2.6 其他正交变换



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

- 快速傅立叶变换的算法就是在研究离散傅立叶变换计算的基础上，节省计算量以达到快速计算的目的。由于二维离散傅立叶变换可以分离成两步的一维离散傅立叶变换来实现，因此这里只介绍一维离散傅立叶变换的快速算法。
- 快速傅立叶变换的设计思想是将原函数分类为奇数项和偶数项，通过两项不断地进行运算最终得到需要的结果。





第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ 矩阵方程：考虑离散傅立叶变换

$$X(n) = \sum_{k=0}^{N-1} x_0(k) e^{-j2\pi nk/N}, \quad n = 0, 1, \dots, N-1$$

上面的式子代表了 N 个方程的计算，为方便表示，我们引入下面一个记号： $W = e^{-j2\pi/N}$ 。则：

$$\sum_{k=0}^{N-1} x_0(k) W^k$$

如果 $N = 4$ ，则方程可写为：

$$\begin{cases} X(0) = x_0(0)W^0 + x_0(1)W^0 + x_0(2)W^0 + x_0(3)W^0 \\ X(1) = x_0(0)W^0 + x_0(1)W^1 + x_0(2)W^2 + x_0(3)W^3 \\ X(2) = x_0(0)W^0 + x_0(1)W^2 + x_0(2)W^4 + x_0(3)W^6 \\ X(3) = x_0(0)W^0 + x_0(1)W^3 + x_0(2)W^6 + x_0(3)W^9 \end{cases}$$





第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ 矩阵表示：

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ X(3) \end{bmatrix} = \begin{bmatrix} W^0 & W^0 & W^0 & W^0 \\ W^0 & W^1 & W^2 & W^3 \\ W^0 & W^2 & W^4 & W^6 \\ W^0 & W^3 & W^6 & W^9 \end{bmatrix} \begin{bmatrix} x_0(0) \\ x_0(1) \\ x_0(2) \\ x_0(3) \end{bmatrix}$$

或者表示成：

$$X(n) = W^k x_0(k)$$

□ 矩阵的计算次数：要完成矩阵的运算，需要做 $N \times N$ 次复数的乘法和 $N(N - 1)$ 次复数加法



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ W_N^{ux} 的性质

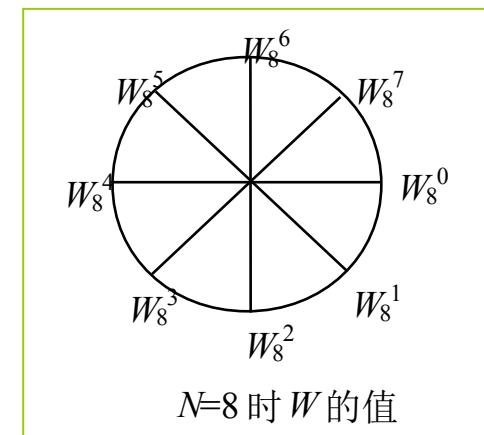
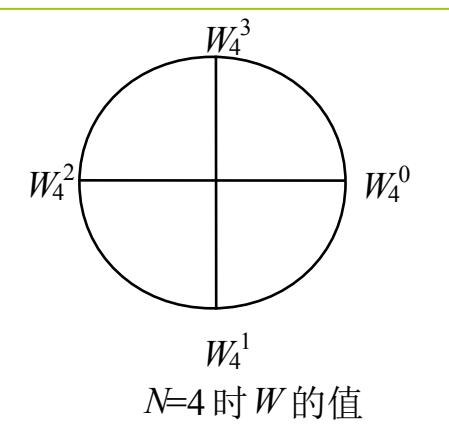
◆ 对称性

$$W_N^{u(N-x)} = W_N^{-ux} = (w_N^{ux})^*$$

◆ 周期性

$$W_N^{ux} = W_N^{(u+N)x} = w_N^{u(N+x)}$$

单位圆表示





第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ 改写矩阵

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ X(3) \end{bmatrix} = \begin{bmatrix} W^0 & W^0 & W^0 & W^0 \\ W^0 & W^1 & W^2 & W^3 \\ W^0 & W^2 & W^4 & W^6 \\ W^0 & W^3 & W^6 & W^9 \end{bmatrix} \begin{bmatrix} x_0(0) \\ x_0(1) \\ x_0(2) \\ x_0(3) \end{bmatrix}$$

这是因为：

$$W^{nk} = W^{(nk)\bmod(N)}$$

例如， $N = 4, n = 2, k = 3$ ，则：

$$W^2 = e^{-j2\pi*2/4} = e^{-j\pi}$$



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ 矩阵分解因子：

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ X(3) \end{bmatrix} = \begin{bmatrix} 1 & W^0 & 0 & 0 \\ 1 & W^2 & 0 & 0 \\ 0 & 0 & 1 & W^1 \\ 0 & 0 & 1 & W^3 \end{bmatrix} \begin{bmatrix} 0 & 1 & W^0 & 0 \\ 0 & 1 & 0 & W^0 \\ 1 & 0 & W^2 & 0 \\ 0 & 1 & 0 & W^2 \end{bmatrix} \begin{bmatrix} x_0(0) \\ x_0(1) \\ x_0(2) \\ x_0(3) \end{bmatrix}$$

注：上面的列矢量 $x(n)$ 的行顺序发生了改变。

□ 乱序后的列矢量：用下面的符号标记乱序后的列矢量

$$\bar{X}(n) = \begin{bmatrix} x(0) \\ x(2) \\ x(1) \\ x(3) \end{bmatrix}$$



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ 计算次数：将矩阵分解因子后，计算需要分两步来进行

◆ 第一步：

$$\begin{bmatrix} x_1(0) \\ x_1(1) \\ x_1(2) \\ x_1(3) \end{bmatrix} = \begin{bmatrix} 1 & 0 & W^0 & 0 \\ 0 & 1 & 0 & W^0 \\ 1 & 0 & W^2 & 0 \\ 0 & 1 & 0 & W^2 \end{bmatrix} \begin{bmatrix} x_0(0) \\ x_0(1) \\ x_0(2) \\ x_0(3) \end{bmatrix}$$

$$x_1(0) = x_0(0) + W^0 x_0(2)$$

$$x_1(1) = x_0(1) + W^0 x_0(3)$$

其中：

$$x_1(2) = x_0(0) + W^2 x_0(2) = x_0(0) - W^0 x_0(2)$$

$$x_1(3) = x_0(1) + W^2 x_0(3) = x_0(1) - W^0 x_0(3)$$



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ 计算次数：(第二步)

$$\begin{bmatrix} x(0) \\ x(2) \\ x(1) \\ x(3) \end{bmatrix} = \begin{bmatrix} x_2(0) \\ x_2(1) \\ x_2(2) \\ x_2(3) \end{bmatrix} = \begin{bmatrix} 1 & W^0 & 0 & 0 \\ 1 & W^2 & 0 & 0 \\ 0 & 0 & 1 & W^1 \\ 0 & 0 & 1 & W^3 \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_1(1) \\ x_1(2) \\ x_1(3) \end{bmatrix}$$

$$\text{其中: } x_2(0) = x_1(0) + W^0 x_1(1)$$

$$x_2(2) = x_1(2) + W^1 x_1(3)$$

$$\text{而: } x_2(1) = x_1(0) - W^0 x_1(1)$$

$$x_2(3) = x_1(2) - W^1 x_1(3)$$



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ **计算次数：** 经过矩阵分解后，计算方程总共需要四次复数乘法和八次复数加法。而未经分解的矩阵计算，总共

需要十六次复数乘法和十二次复数加法。

□ **效率：** 因为计算时间主要取决于复数乘法的计算次数，所以减少复数乘法的次数就是 FFT 算法效率高的原因

○ □ **基2的 FFT 算法的原理：** 对于 $N = 2^\gamma$ 的 FFT 算法，就是要把一个 $N \times N$ 的矩阵，分解为 γ (其中每个都是 $N \times N$) 个矩阵。使被分解后的每一个矩阵具有复数乘法和复数加法最少的特性。





第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ 基2的FFT算法的效率：对于 $N = 2^\gamma$ 的FFT算法，所需的计算次数为：

◆ 乘法次数： $\frac{N \cdot \gamma}{2}$

◆ 加法次数： $N \cdot \gamma$

□ 普通算法的效率：所需的计算次数为：

◆ 乘法次数： $N \cdot N$

◆ 加法次数： $N(N - 1)$

□ 两种算法的效率之比：

$$\frac{N^2}{N \cdot \gamma / 2} = \frac{2N}{\gamma}$$



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

- 1965 年， Cooley 和 Tukey 提出快速算法，算法时间复杂度为 $N \log_2 N$ 。

N	DFT 计算量	FFT 计算量	比值
8	64	24	2.7
16	256	64	4.0
32	1024	160	6.4
64	4096	384	10.7
128	16384	896	18.3
256	65536	2048	32.0
512	262144	4608	56.9
1024	1048576	10240	102.4
2048	4194304	22528	186.2



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ 乱序重排：经过矩阵分解后，计算所得到的是一个乱序的列矢量，这种乱序是分解过程中固有的，需要经过重新排列。

◆ 二进制表示：将列矢量的自变量表示成二进制

$$\bar{X}(n) = \begin{bmatrix} x(0) \\ x(2) \\ x(1) \\ x(3) \end{bmatrix} = \begin{bmatrix} x(00) \\ x(10) \\ x(01) \\ x(11) \end{bmatrix}$$

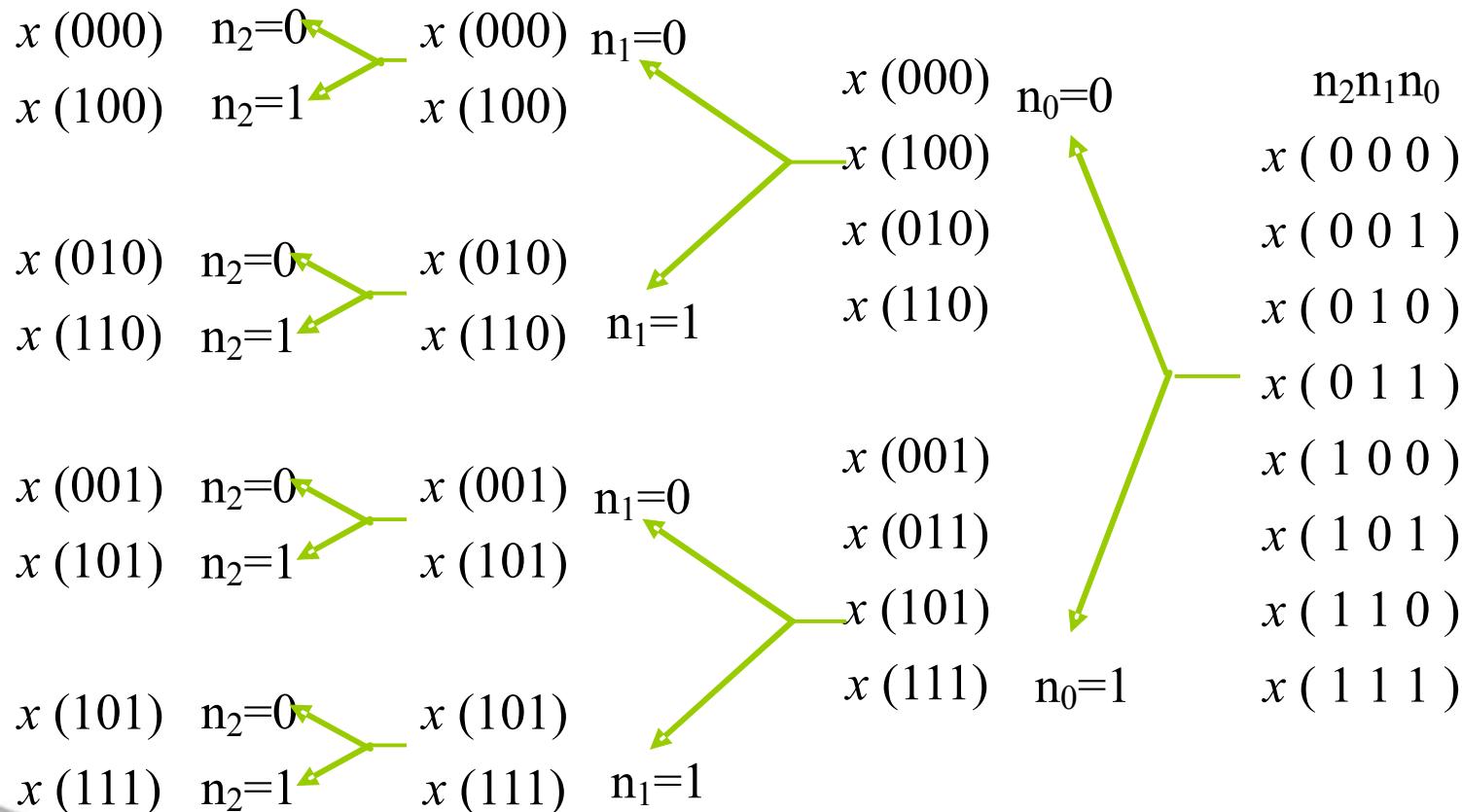
◆ 位序颠倒：将列矢量的自变量二进制码的位序颠倒

$$\begin{bmatrix} x(00) \\ x(10) \\ x(01) \\ x(11) \end{bmatrix} \text{ 翻转成 } \begin{bmatrix} x(00) \\ x(01) \\ x(10) \\ x(11) \end{bmatrix} = x(n)$$



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换分组和比特倒序

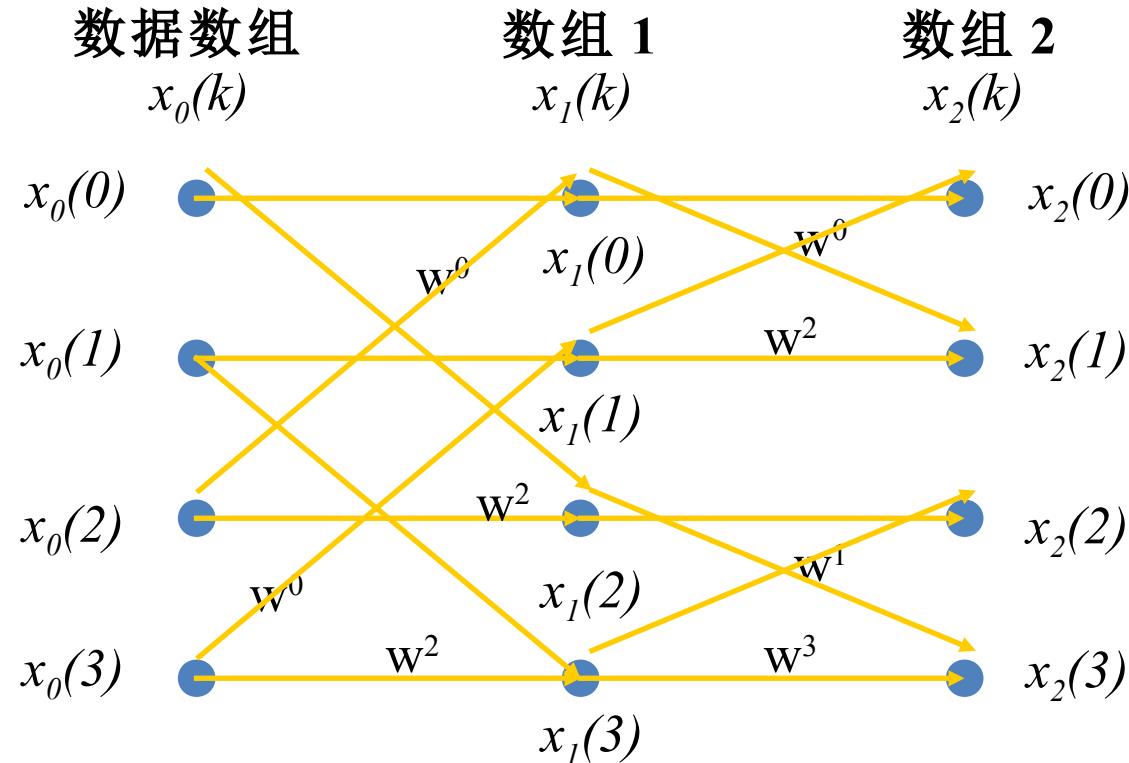




第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ 信号流程图：前面矩阵分解计算的过程可以用下面的信号流程图表示出来：





第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

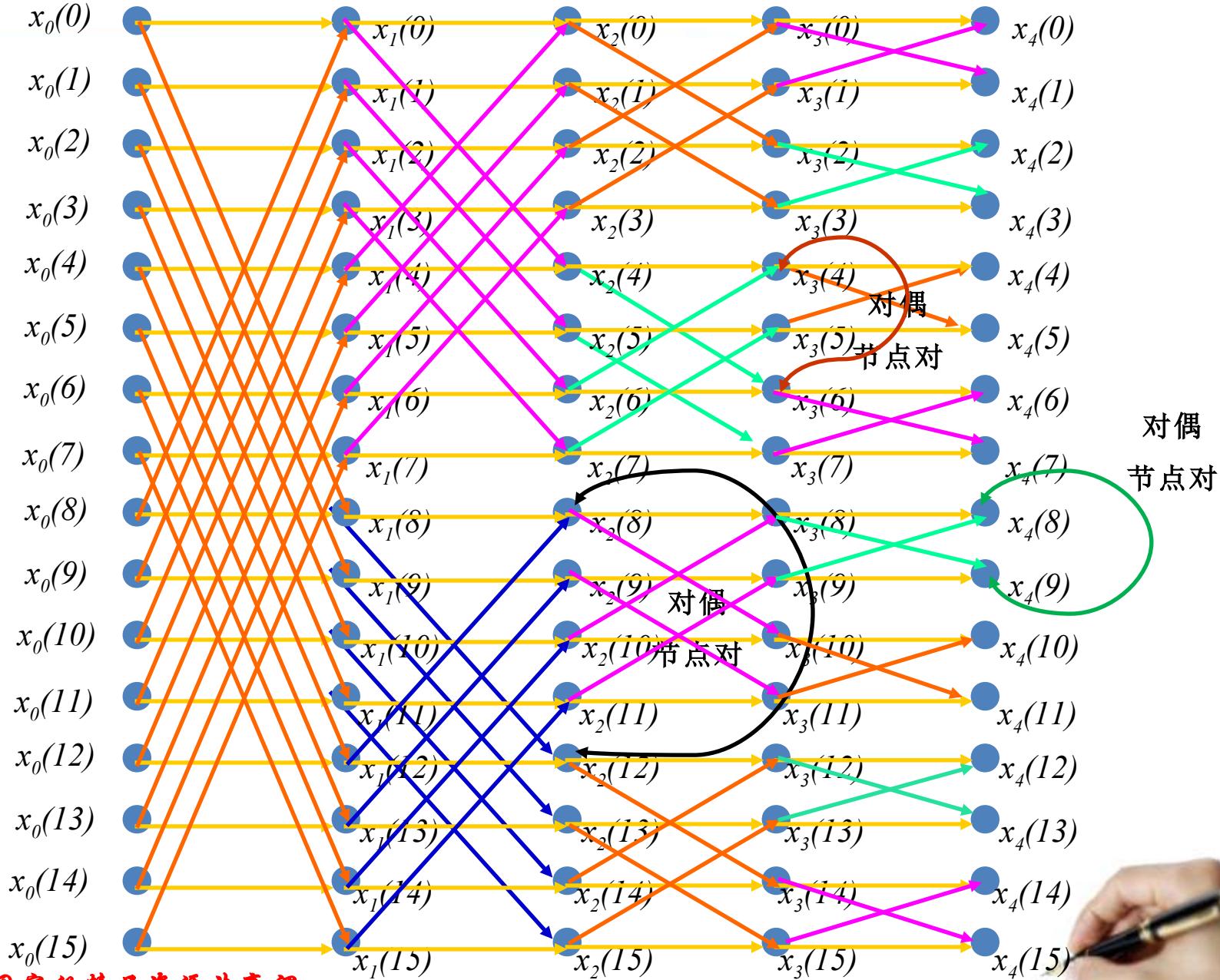
对信号流程图的几点说明：

- ◆ **传输路径：**进入计算数组的每个节点有两条实线，它们表示从上一列节点来的两条传输路径。
- ◆ **传输路径的权值：**每条传输路径都带有相应的权值。如果在某条传输路径上没有标记权值，则缺省权值为1。
- ◆ **数组的计算：**从两条传输路径进到一个节点的两结果要相加起来。



数据数组

对偶节点对





第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ 对偶节点：在前面的图中的每一纵列中，总可以找到这样的两个点，

它们的传输路径来自前一列中的同一对节点，而且这两个节点不会用来计算其它的任何节点。我们把这样的两个点称为对偶节点。例

如， $x_1(0)$ 和 $x_1(8)$ ，它们都用 $x_0(0)$ 和 $x_0(8)$ 来计算。可以用 $x_0(0)$ 和 $x_0(8)$ 同时计算出 $x_1(0)$ 和 $x_1(8)$ 。

□ 对偶节点的间隔：用 l 表示数组的下标，在 $l=1$ 的列中，对偶节点（例如 $x_1(0)$ 和 $x_1(8)$ ）之间的距离为 $K=8=N/2^1$ ，在 $l=2$ 的列中，对偶节点（例如 $x_2(8)$ 和 $x_2(12)$ ）之间的距离为 $K=4=N/2^2$ ，同样，在 $l=3$ 的列中，对偶节点之间的距离为 $K=2=N/2^3$ ，在 $l=4$ 的列中，对偶节点之间的距离为 $K=1=N/2^4$ 。所以，对偶节点的间隔为：

$$K = \frac{N}{2^l}$$



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

- 对偶节点对的计算次数：一个对偶节点对的计算只需要一次复数乘法。一般地讲，如果在某一节点上的加权系数为 W^p ，则在其对偶节点上的加权系数就是 $W^{p+N/2}$ ，因为 $W^p = W^{p+N/2}$ ，所以计算对偶节点对，只需一次复数乘法。
- 对偶节点对的计算公式：任何一个对偶节点对的计算，都可以用下面的一对公式来给定：

$$\begin{cases} x_l(k) = x_{l-1}(k) + W^p x_{l-1}\left(k + \frac{N}{2^l}\right) \\ x_l\left(k + \frac{N}{2^l}\right) = x_{l-1}(k) - W^p x_{l-1}\left(k + \frac{N}{2^l}\right) \end{cases}$$



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ W^p 的确定：每个节点的加权系数可以通过下面的方法得到

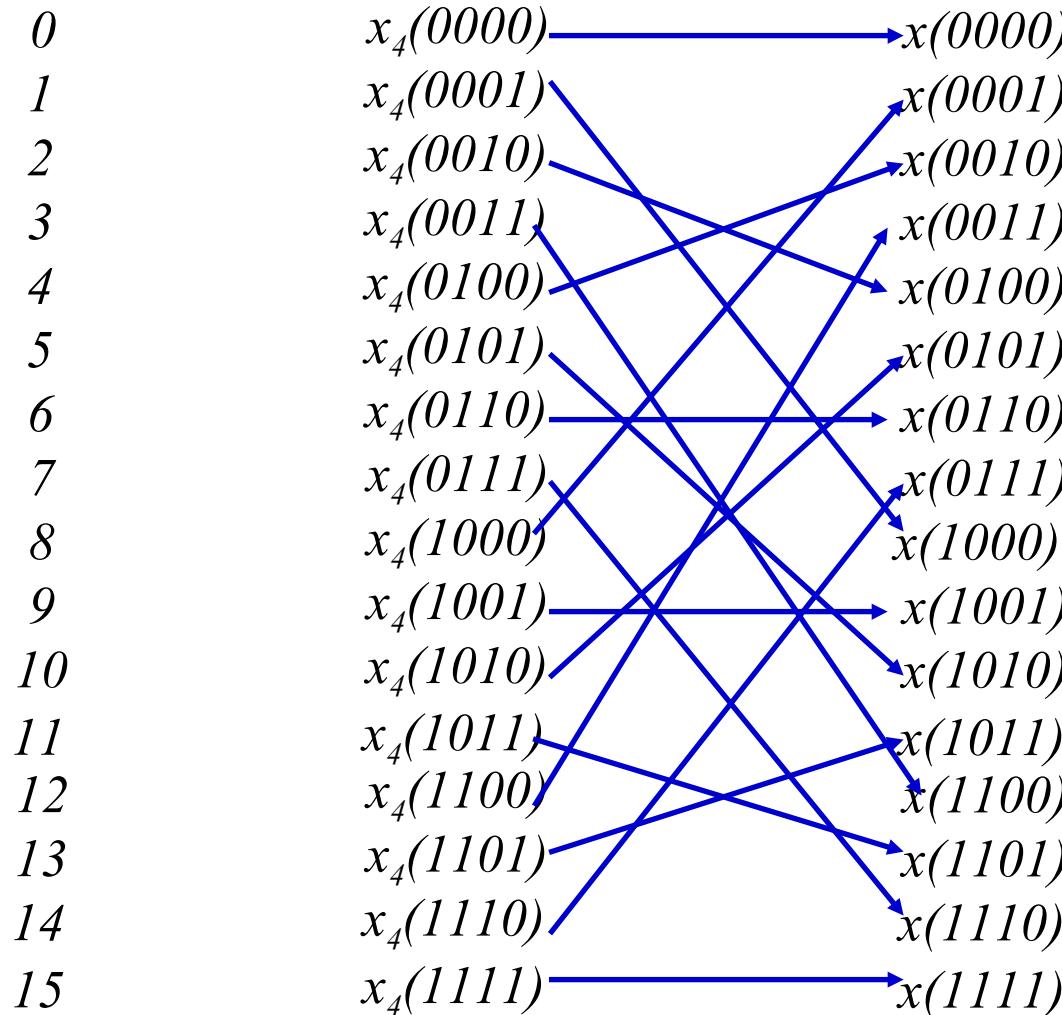
- ◆ 把指数 k 写成 γ 位的二进制数；
- ◆ 把这个二进制数右移 $\gamma-1$ 位，并把左边空出的位置补为零；
- ◆ 将右移后的 γ 位二进制数的位序颠倒，其结果就是 P 值；
- ◆ 例如，前面图中的节点 $x_3(8)$ ，由于 $\gamma=4$ ， $k=8$ ， $l=3$ ，于是 k 写成 γ 位的二进制数就是 1000，将这个二进制数右移 $\gamma-1 = 4-3=1$ 位，并将左边空出的位置上补零，结果是 0100，然后把位序颠倒。得到 0010，这就是十进制的整数 2，所以 p 值等于 2。





第二章 常用的数学变换

- **FFT 的整序：** 整序的方法在前面已经提到过，就是将最后得到的计算数组的指数 n 先写成二进制形式，再进行位序颠倒。

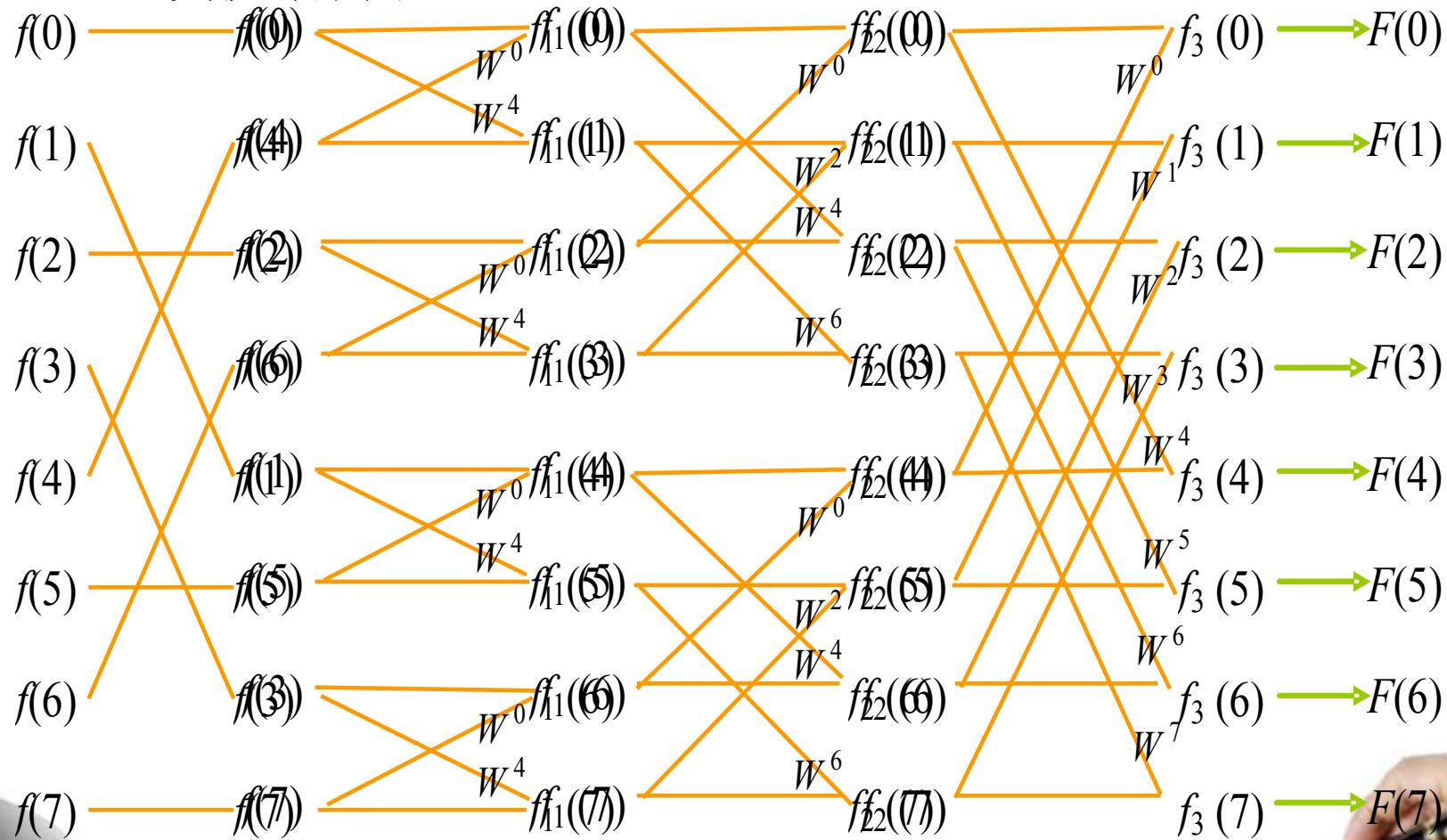




第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

FFT 变换蝶形图：





第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ FFT计算实例

设 有函数 $f(x)$, 其 $N = 2^3 = 8$

则有 $\{f(0), f(1), f(2), f(3), f(4), f(5), f(6), f(7)\}$

计算 $\{F(0), F(1), F(2), F(3), F(4), F(5), F(6), F(7)\}$



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换

□ FFT计算实例

◆ 首先分成奇偶两组：

奇： $\{f(0), f(2), f(4), f(6)\}$

偶： $\{f(1), f(3), f(5), f(7)\}$

◆ 为了利用传递性，再分成两组：

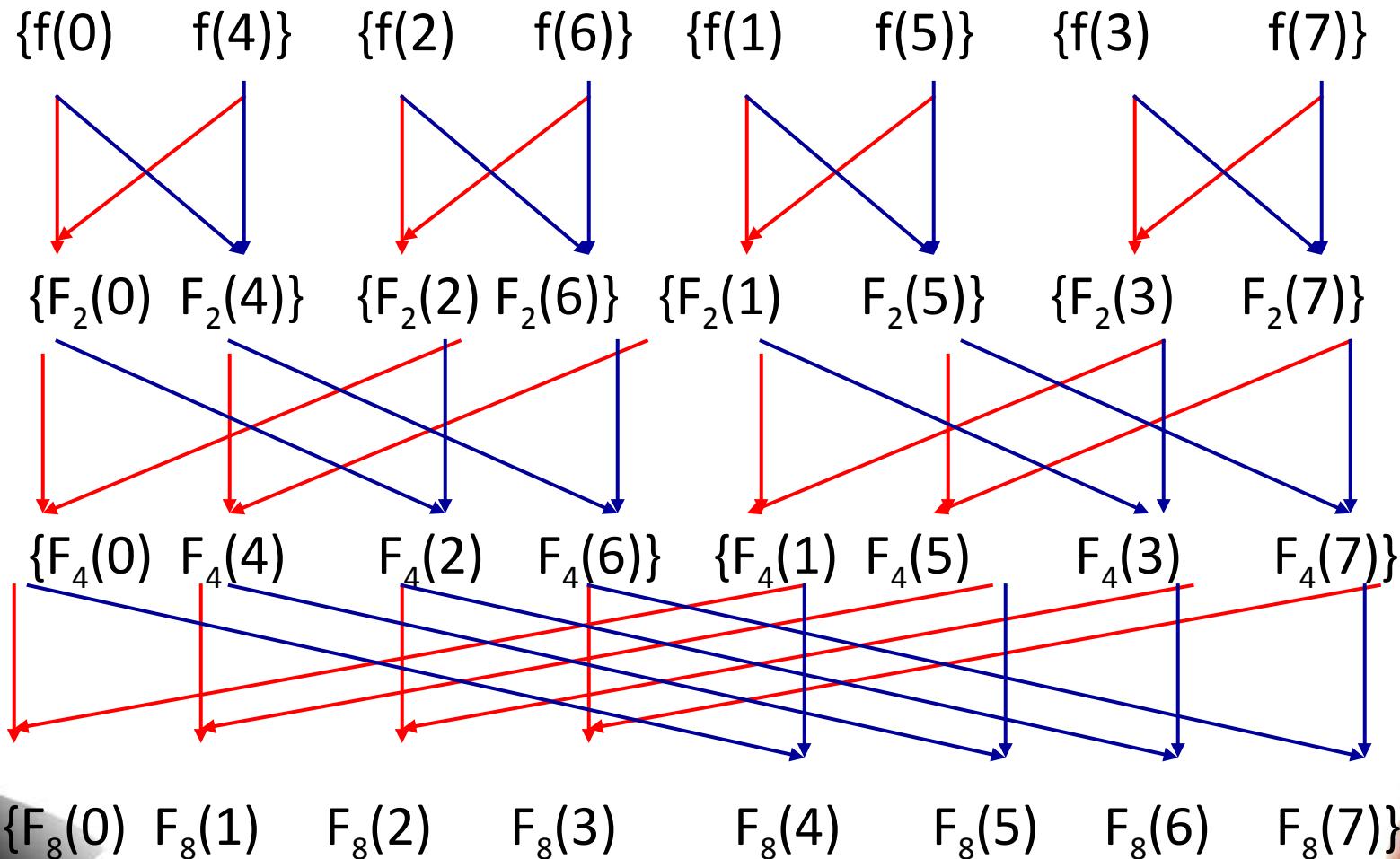
$\{f(0), f(4)\}, \{f(2), f(6)\}$

$\{f(1), f(5)\}, \{f(3), f(7)\}$



第二章 常用的数学变换

2.3 频率域变换（傅立叶变换）——2.3.3 快速傅立叶变换





国家级精品资源共享课

数字图像处理

(Digital Image Processing)

山东科技大学
江涛教授



第二章 常用的数学变换

- 2.1 引言
- 2.2 空域变换
- 2.3 频率域变换
- 2.4 离散余弦变换
- 2.5 KL 变换
- 2.6 其他正交变换



第二章 常用的数学变换

2.4 离散余弦变换 (DCT)

- 傅立叶变换的最大问题是：它的参数都是复数，在数据的描述上相当于实数的2倍。为此，我们希望有一种能够达到相同功能但数据量又不大的变换。

- 当 $f(x)$ 或 $f(x, y)$ 为偶函数时，傅立叶变换的计算公式虚部为0，只有余弦项。

- 余弦变换是简化傅立叶变换的一种方法。



第二章 常用的数学变换

2.4 离散余弦变换

- 一维离散余弦变换定义

$$F(u) = \sqrt{\frac{2}{N}} \sum_{x=0}^{N-1} f(x) \cos \left[\frac{\pi}{2N} (2x + 1)u \right]$$

$$F(0) = \sqrt{\frac{2}{N}} \sum_{x=0}^{N-1} f(x)$$

通过归一化表示为：

$$F(u) = c(u) \sqrt{\frac{2}{N}} \sum_{x=0}^{N-1} f(x) \cos \left[\frac{\pi}{2N} (2x + 1)u \right]$$

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



第二章 常用的数学变换

2.4 离散余弦变换

□ 一维离散余弦变换定义

$$F(u) = \sqrt{\frac{2}{N}} \sum_{x=0}^{N-1} c(u) f(x) \cos \left[\frac{\pi}{2N} (2x+1) u \right]$$

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

矩阵表示:

$$F = Cf, f = C^T F$$



第二章 常用的数学变换

2.4 离散余弦变换 (DCT)

将一维离散余弦变换扩展到二维离散余弦变换，设 $f(x, y)$ 为 $M \times N$ 的数字图像矩阵，则：

$$F(u, v) = \frac{2}{\sqrt{MN}} C(u)C(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cos \frac{(2x+1)u\pi}{2M} \cos \frac{(2y+1)v\pi}{2N}$$

二维 DCT 逆变换定义如下：

$$f(x, y) = \frac{2}{\sqrt{MN}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} C(u)C(v) F(u, v) \cos \frac{(2x+1)u\pi}{2M} \cos \frac{(2y+1)v\pi}{2N}$$

式中： $x, u=0, 1, 2, \dots, M-1$ ； $y, v=0, 1, 2, \dots, N-1$



第二章 常用的数学变换

2.4 离散余弦变换

□ 若 $M = N$ ，则：

$$Q(u) = Q(v) = \begin{cases} \frac{1}{\sqrt{2}} & u, v = 0 \\ 1 & \text{其它} \end{cases}$$

类似一维矩阵形式的DCT，可以写出二维DCT的矩阵形式如下：

$$F_c = CfC, \quad f = CF_cC$$

其中，矩阵 C 的元素为：

$$c_{i,m} = \alpha(m) \cos \left[\frac{\pi(2i+1)m}{2N} \right], \quad \alpha(m) = \begin{cases} \sqrt{\frac{1}{N}} & m = 0 \\ \sqrt{\frac{2}{N}} & m \neq 0 \end{cases}$$



第二章 常用的数学变换

2.4 离散余弦变换

□ 正变换矩阵为：

$$\sqrt{\frac{2}{N}} \begin{bmatrix} 1 & \frac{1}{\sqrt{2}} & \cdots & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \cos \frac{\pi}{2N} & \cdots & \cos \frac{(2N-1)\pi}{2N} \\ \cos \frac{\pi}{2N} & \cos \frac{3\pi}{2N} & \cdots & \cos \frac{(2N-1)(N-1)\pi}{2N} \\ \vdots & \vdots & & \vdots \\ \cos \frac{(N-1)\pi}{2N} & \cos \frac{3(N-1)\pi}{2N} & \cdots & \cos \frac{(2N-1)(N-1)\pi}{2N} \end{bmatrix}$$

□ $N=4$ 时：

$$\begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 \\ 0.653 & 0.271 & -0.271 & -0.653 \\ 0.5 & -0.5 & -0.5 & 0.5 \\ 0.271 & -0.653 & 0.653 & -0.271 \end{bmatrix}$$



第二章 常用的数学变换

2.4 离散余弦变换

(DCT)

□ 余弦变换的性质

- ◆ 余弦变换为实的正交变换，变换核的基函数正交
- ◆ 序列的余弦变换是 DFT 的对称扩展形式
- ◆ 核可分离，可以用两次一维变换来执行
- ◆ 余弦变换的能量向低频集中
- ◆ 余弦变换有快速变换，和傅立叶变换一样，分奇偶组



第二章 常用的数学变换

2.4 离散余弦变换 (DCT)

□ 通常根据可分离性，二维 DCT 可用两次一维 DCT 来完成，其算法流程与 DFT 类似，即：

$$f(x, y) \rightarrow F_{\text{行}}[f(x, y)] = F(x, v)$$

转置

$$\rightarrow F(x, v)^T \rightarrow F_{\text{列}}[F(x, v)^T] = F(u, v)^T$$

转置

$$\rightarrow F(u, v)$$



第二章 常用的数学变换

2.4 离散余弦变换

- (DCT)
- DCT示例

已知一个矩阵

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

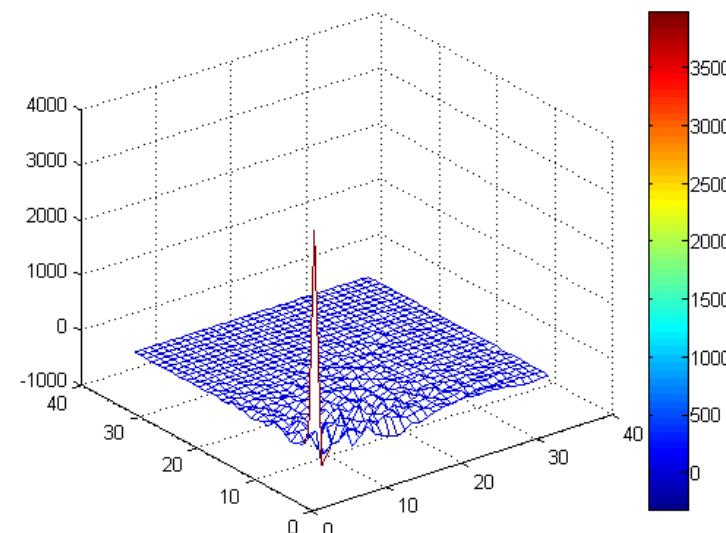
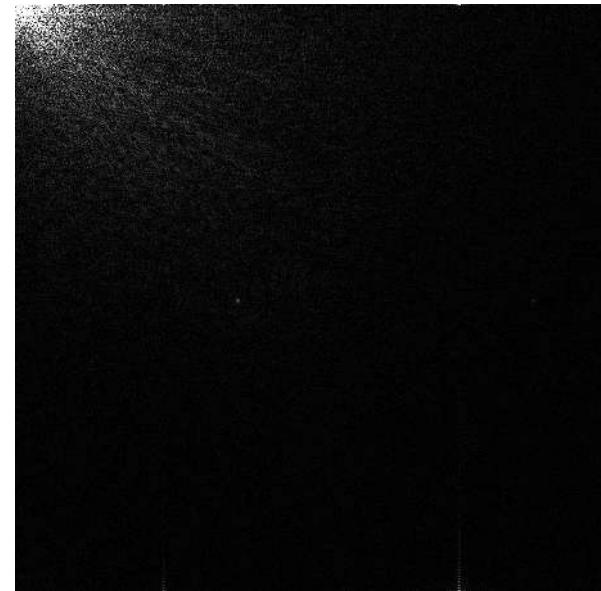
求DCT变换

$$A_{DCT} = \begin{bmatrix} 5 & -1 \\ -2 & 0 \end{bmatrix}$$



第二章 常用的数学变换

2.4 离散余弦变换 (DCT)



其变换结果的能量将主要集中在左上角的位置，即低频部分



第二章 常用的数学变换

2.4 离散余弦变换

	1	(DCT)	3	4	5	6	7	8
1	137	136	133	136	138	134	134	132
2	137	136	133	136	138	134	134	132
3	138	133	134	134	136	132	130	130
4	133	133	133	130	134	133	128	125
5	129	133	130	130	133	131	132	128
6	131	133	130	122	132	131	130	130
7	131	130	130	130	132	131	128	130
8	131	132	130	130	131	131	130	128
9	129	130	130	127	130	129	130	129
10	129	130	132	129	131	129	132	134
11	130	130	129	128	131	130	133	132
12	131	129	125	130	132	130	131	132

Lena 原图矩阵

	1	2	3	4	5	6	7	8
1	25262	-3291.4	-473.31	2518.3	2102.7	-1164.9	-167.57	1092.9
2	1526.6	1768.4	-2110.1	-452.08	-99.837	842.71	19.981	-592.77
3	-212.04	1089.5	-1408	2281.2	-1573.4	2706.8	45.043	-315.39
4	-617.46	-407.78	2378.4	-1882.1	294.47	991.15	-1715.2	796.72
5	-884.86	684.19	688.86	580.37	-140.31	471.06	-85.693	274.32
6	-51.517	-186.2	1294.7	297.75	-1301.6	-349.83	949.82	1458.8
7	715.79	-640.97	879.99	-493.09	-580.74	-48.366	1153.4	-1027.3
8	982.43	-693.21	-450.9	729.31	-354.86	397.48	396.2	-805.73
9	-65.387	56.909	-23.731	-129.43	321.4	-256.92	-475.82	-181.47
10	137.57	-419.25	36.447	-115.45	534.56	-704.79	17.464	409.75
11	-304.12	339.01	-342.76	0.19983	471.95	-847.56	342.1	457.45
12	-495.32	159.9	341.2	-571.27	72.808	-243.06	9.8179	-253.91

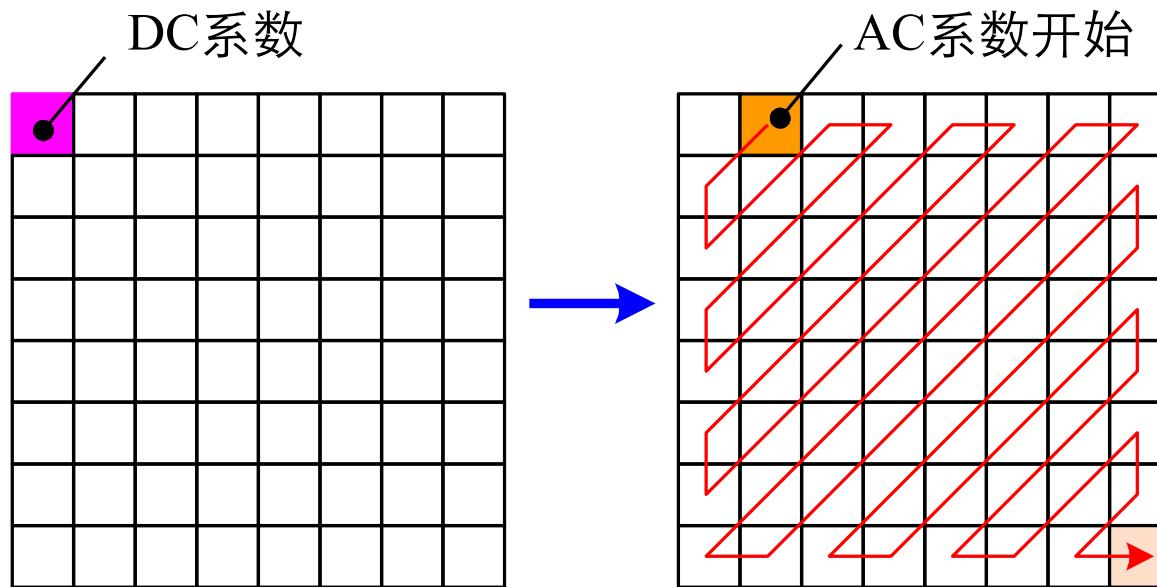
Lena DCT 系数矩阵



第二章 常用的数学变换

2.4 离散余弦变换 (DCT)

“Z”（“之”）形编码排序





第二章 常用的数学变换

2.4 离散余弦变换

(DCT)

采样数据(压缩数据)

89	101	114	125	126	115	105	96
97	115	131	147	149	135	123	113
114	134	159	178	175	164	149	137
121	143	177	196	201	189	165	150
119	141	175	201	207	186	162	144
107	130	165	189	192	171	144	125
97	119	149	171	172	145	117	96
88	107	136	156	155	129	97	75

DCT

1125	-32	-185	-7	2	-1	-2	2
-22	-16	45	-3	-2	0	-2	-2
-165	32	17	9	5	1	-3	0
-7	-4	0	2	2	-1	-1	2
-2	0	0	3	0	0	2	1
3	1	1	-1	-2	0	2	0
0	0	2	-1	-1	2	1	-1
0	3	1	-1	2	1	-2	0

量化
(△:4)

281	-8	-46	-1	0	0	0	0
-5	-4	11	0	0	0	0	0
-41	8	4	2	1	0	0	0
-1	-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

“之”字型读出

编码

281,-8,-5,-41,-4,-46,-1,11,8,-1
0,-1,4,0,0,0,2,0,0,0,0,0,0
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
281,-8,-5,-41,-4,-46,-1,11,8,-1
1'0,-1,4,4'0,2,7'0,1,EOB



第二章 常用的数学变换

- 2.1 引言
- 2.2 空域变换
- 2.3 频率域变换
- 2.4 离散余弦变换
- 2.5 K-L 变换**
- 2.6 其他正交变换



第二章 常用的数学变换

2.5 K-L 变换

- K-L 变换，也称为 Hotelling 变换、特征向量变换 (Eigenvector-Based Transform) 、主成分分析 (PCA Principle Component Analysis) 等。
- KL 变换方法是应用最广泛的一种特征提取方法之一，它是一种利用图像的统计性质的变换。在信号处理、模式识别、数字图像处理等领域已经得到了广泛的应用。其基本思想是提取出空间原始数据中的主要特征，减少数据冗余，使得数据在一个低维的特征空间被处理，同时保持原始数据的绝大部分的信息，从而解决数据空间维数过高的瓶颈问题。常用在数据压缩、特征提取等方面。

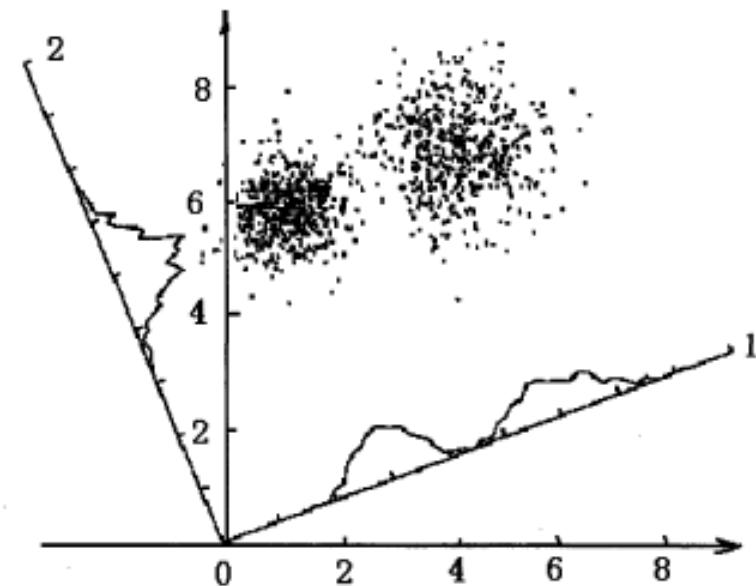




第二章 常用的数学变换

2.5 K-L 变换

□ 图中水平轴和垂直轴表示数据集的自然坐标轴。标号为 1 和 2 的旋转坐标轴是应用这个数据集的主变量分析产生的结果。从图可以看出数据集投影到 1 号轴上抓住了数据的主要特征，即具有双峰（即在它的结构上有两个聚类）的特点。



双变量(二维)数据集



第二章 常用的数学变换

2.5 K-L 变换

□ K-L 变换原理

设 $X = [x_1, x_2, \dots, x_n]^T$ 和 $Y = [y_1, y_2, \dots, y_n]^T$

为两个 n 维随即向量，其元素 x_i, y_i 分别具有 M 个随机值假定 X 能由 Y 精确表示为：

$$X = \Phi Y$$

Φ 为 $n \times n$ 正交矩阵，记为 $\Phi = [\phi_1, \phi_2, \dots, \phi_n]$

若取 Y 向量的前 m 个向量 Y_m 来表示 X ，记为 X_m ，可有误差

$$\Delta X_m = X - X_m$$

从统计角度，如何选择 Φ ，使得上述误差的统计均方值达到极小



第二章 常用的数学变换

2.5 K-L 变换

□ K-L变换原理

- ◆ 通过数学分析，可得出结论如下：

对正交矩阵 $\Phi = [\phi_1, \phi_2, \dots, \phi_n]$

- ◆ 若取 ϕ_i 为 X 的协方差矩阵 C_x 的特征向量，则对 X 进行下述变换后

$$Y = \Phi^T X$$

其结果 Y 向量可满足前述要求

- ◆ 上述变量式与反变换

$$X = \Phi^T Y$$



第二章 常用的数学变换

2.5 K-L 变换

□ K-L变换计算

- ◆ 对 N 维随机向量 $X = [x_1, x_2, \dots, x_n]^T$, 其每个元素 x_i 分别具有 M 个样本
- ◆ 其平均值向量定义为: $M_x = E\{X\}, M_x \cong \frac{1}{M} \sum_{i=1}^M x_i$
- ◆ 其协方差矩阵为一个 $N \times N$ 的矩阵, 定义为:

$$C_x = E\{(X - M_x)(X - M_x)^T\}$$

$$C_x \cong \frac{1}{M} \sum_{i=1}^M (x_i - m_x) (x_i - m_x)^T = \frac{1}{M} \left[\sum x_i x_i^T \right] - m_x m_x^T$$



第二章 常用的数学变换

2.5 K-L 变换

□ K-L变换计算

- ◆ 令 F 和 λ 为 C_x 的特征向量和对应的特征值，可有关系

$$|C_x - \lambda I| = 0, C_x F = \lambda F$$

- ◆ 特征向量 F 为 N 维矢量，由上式可解出 N 个特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ ，将其按降序排列 $\lambda_1 > \lambda_2 > \lambda_3 > \dots > \lambda_n$
- ◆ 将各特征值分别代入上式，可得出对应各特征值的特征向量为：

$$F_i = [f_{i1}, f_{i2}, \dots, f_{in}]^T$$



第二章 常用的数学变换

2.5 K-L 变换

□ K-L变换计算

- ◆ 将各特征向量转置后即可构成变换矩阵:

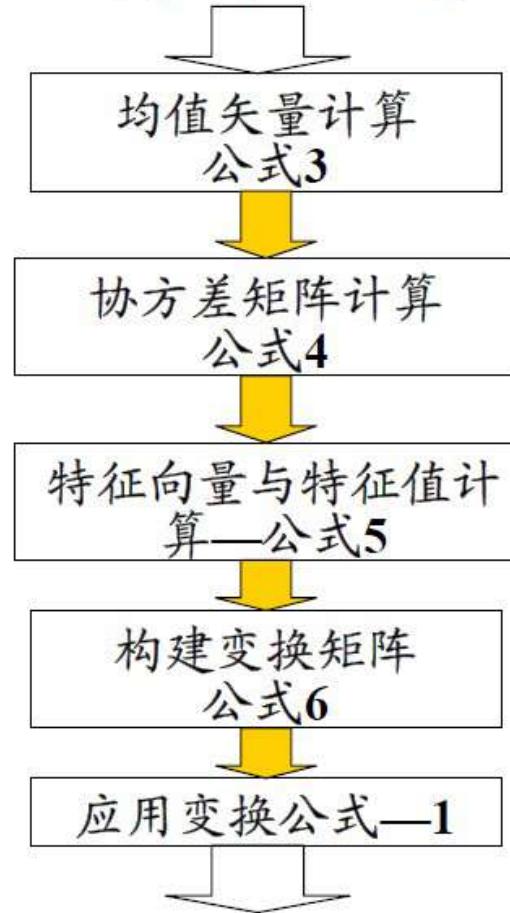
$$\Phi = \begin{bmatrix} F_1^T \\ F_2^T \\ \vdots \\ F_N^T \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1N} \\ f_{21} & f_{22} & \cdots & f_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ f_{N1} & \cdots & \cdots & f_{NN} \end{bmatrix}$$



第二章 常用的数学变换

2.5 K-L 变换

确定随机向量集合
 $X = [x_1, x_2, \dots, x_n]^T$



获取变换结果 $Y = [y_1, y_2, \dots, y_n]^T$





第二章 常用的数学变换

2.5 K-L 变换

□ K-L变换计算实例

- ◆ 假定 2×6 图像矩阵为: $X = \begin{bmatrix} 2 & 4 & 5 & 5 & 3 & 2 \\ 2 & 3 & 4 & 5 & 4 & 3 \end{bmatrix}$
- ◆ 特将每一行作为一个随机矢量, 形成随机向量为:

$$X = [x_1 \quad x_2]$$

- ◆ 其均值矢量为

$$M = \frac{1}{6} \begin{bmatrix} 2 + 4 + 5 + 5 + 3 + 2 \\ 2 + 3 + 4 + 5 + 4 + 3 \end{bmatrix} = \begin{bmatrix} 3.5 \\ 3.5 \end{bmatrix}$$



第二章 常用的数学变

换

2.5 K-L 变换

□ K-L 变换计算实例

◆ 可有协方差矩阵：

$$\begin{aligned} C_x &\cong \frac{1}{6} \sum_{i=1}^6 (x_i - m_x)(x_i - m_x)^T \\ &= \frac{1}{M} \left\{ \begin{bmatrix} 2.25 & 2.25 \\ 2.25 & 2.25 \end{bmatrix} + \begin{bmatrix} 0.25 & -0.25 \\ -0.25 & 0.25 \end{bmatrix} + \begin{bmatrix} 2.25 & 2.25 \\ 2.25 & 2.25 \end{bmatrix} + \begin{bmatrix} 0.25 & -0.25 \\ -0.25 & 0.25 \end{bmatrix} + \begin{bmatrix} 2.25 & 0.75 \\ 0.75 & 0.25 \end{bmatrix} \right\} \\ &= \begin{bmatrix} 1.9 & 1.1 \\ 1.1 & 1.1 \end{bmatrix} \end{aligned}$$



第二章 常用的数学变换

2.5 K-L 变换

□ K-L变换计算实例

◆ 由

$$\begin{vmatrix} 1.9 - \lambda & 1.1 \\ 1.1 & 1.1 - \lambda \end{vmatrix} = 0$$

◆ 可解出: $\lambda^2 - 3\lambda + 0.88 = 0$, $\lambda_1 = 2.67$, $\lambda_2 = 0.33$

◆ 将 λ_1 , λ_2 分别代入公式, 可解出特征向量并构建变换矩阵为:

$$\Phi = \begin{bmatrix} f_1^T \\ f_2^T \end{bmatrix} = \begin{bmatrix} 0.82 & 0.57 \\ -0.57 & 0.82 \end{bmatrix}$$



第二章 常用的数学变换

2.5 K-L 变换

□ K-L变换计算实例

- ◆ 对 X 进行变换

$$Y = \begin{bmatrix} 0.82 & 0.57 \\ -0.57 & 0.82 \end{bmatrix} \begin{bmatrix} 2 & 4 & 5 & 5 & 3 & 2 \\ 2 & 3 & 4 & 5 & 4 & 3 \end{bmatrix}$$
$$= \begin{bmatrix} 2.78 & 4.99 & 6.38 & 6.95 & 4.74 & 3.35 \\ 0.5 & 0.18 & 0.43 & 1.25 & 1.57 & 1.32 \end{bmatrix}$$

- ◆ Y 向量的协方差矩阵为 $C_Y = \begin{bmatrix} 2.67 & 0 \\ 0 & 0.33 \end{bmatrix}$, 可见89%的能量集中在分量1中
- ◆ 舍弃分量2, 反变换结果为:

$$X = \begin{bmatrix} 0.82 & -0.57 \\ 0.57 & 0.82 \end{bmatrix} \begin{bmatrix} 2.78 & 4.99 & 6.38 & 6.95 & 4.74 & 3.35 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 2.28 & 4.1 & 5.23 & 5.7 & 3.89 & 2.75 \\ 1.58 & 2.84 & 3.64 & 3.96 & 2.7 & 1.91 \end{bmatrix}$$





第二章 常用的数学变换

2.5 K-L 变换

□ K-L变换的深入讨论

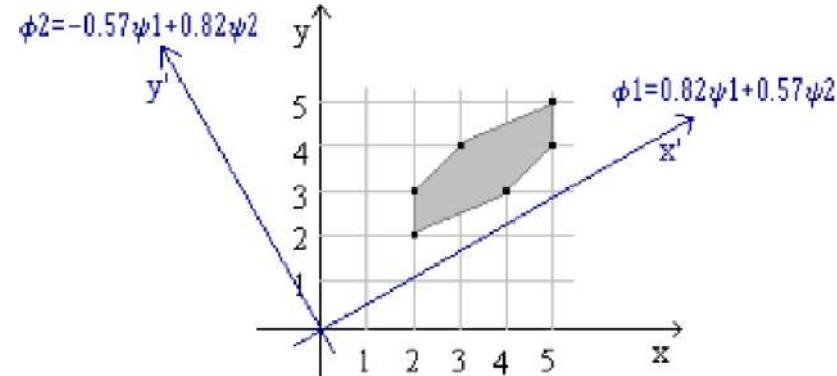
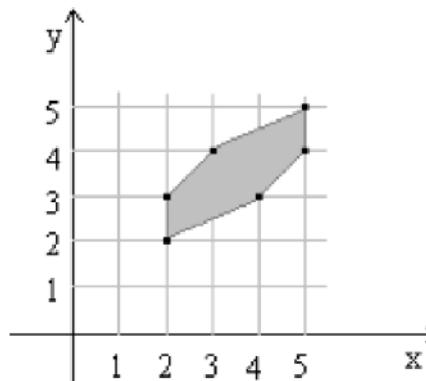
- ◆ 对图像矩阵 X , 可看作为二维平面上一组像素点坐标的集合

$$X = [p_1 \ p_2 \ p_3 \ p_4 \ p_5 \ p_6] = \begin{bmatrix} 2 & 4 & 5 & 5 & 3 & 2 \\ 2 & 3 & 4 & 5 & 4 & 3 \end{bmatrix}$$

- ◆ 变换结果 Y , 则可看作为一个新坐标系下相同像素点集合

$$Y = \begin{bmatrix} 2.78 & 4.99 & 6.38 & 6.95 & 4.74 & 3.35 \\ 0.5 & 0.18 & 0.43 & 1.25 & 1.57 & 1.32 \end{bmatrix}$$

- ◆ 将该新坐标系为原坐标系的旋转, 旋转矩阵即为 $K - L$ 变换矩阵





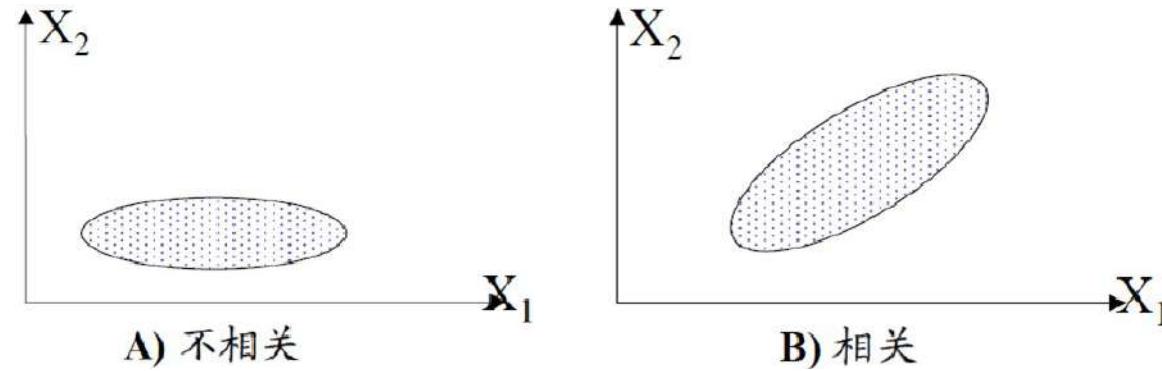
第二章 常用的数学变换

2.5 K-L 变换

□ K-L变换的深入讨论

◆ 考察下图二维 X 空间中框界内各矢量 X :

- 图(B)中, 各点的 x_1 变化时, 其 x_2 也必定增大或减小;
表明矢量 X 的 x_1 , x_2 分量之间**相关性很高**
- 图(A)中, 各点的 x_1 变化时, 其 x_2 变化很小或没有变化;
表明矢量 X 的 x_1 , x_2 分量之间**相关性很低或没有相关性**

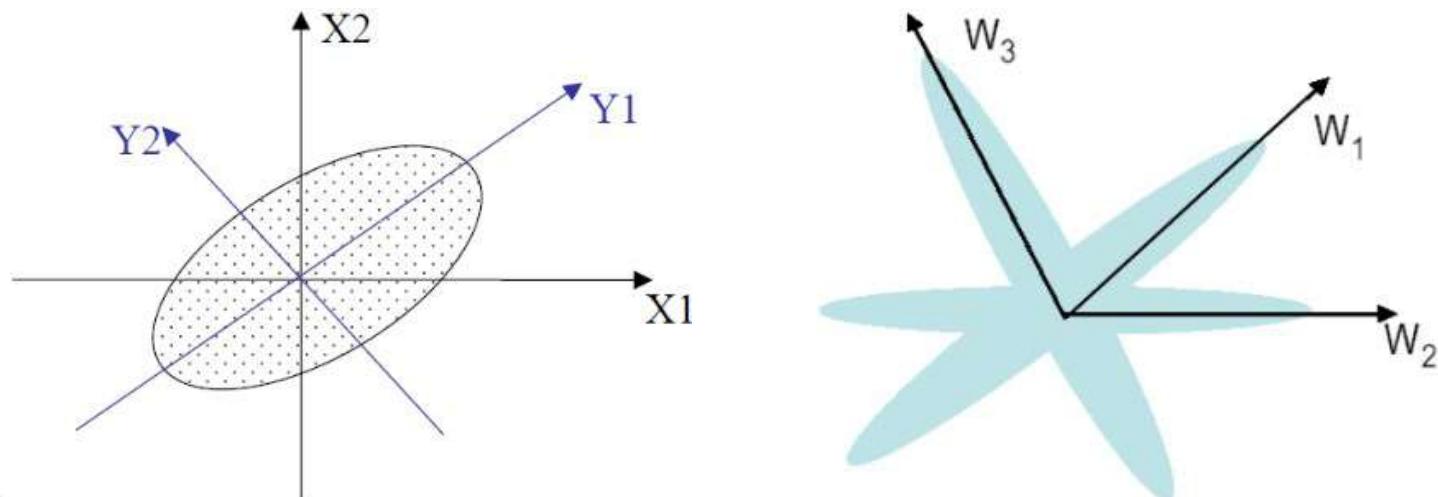




第二章 常用的数学变换

2.5 K-L 变换

- K-L变换的目标，即在于找出使 X 矢量中各分量相关性降低或去除的方向，对图像进行旋转，使其新空间的坐标轴指向各主分量方向-主成分分析或主成分变换
- 扩展至多维空间， $K - L$ 变换可实现多维空间中的去相关，并将能量集中在少数主要分量上





国家级精品资源共享课

数字图像处理

(Digital Image Processing)

山东科技大学
江涛 教授



第二章 常用的数学变换

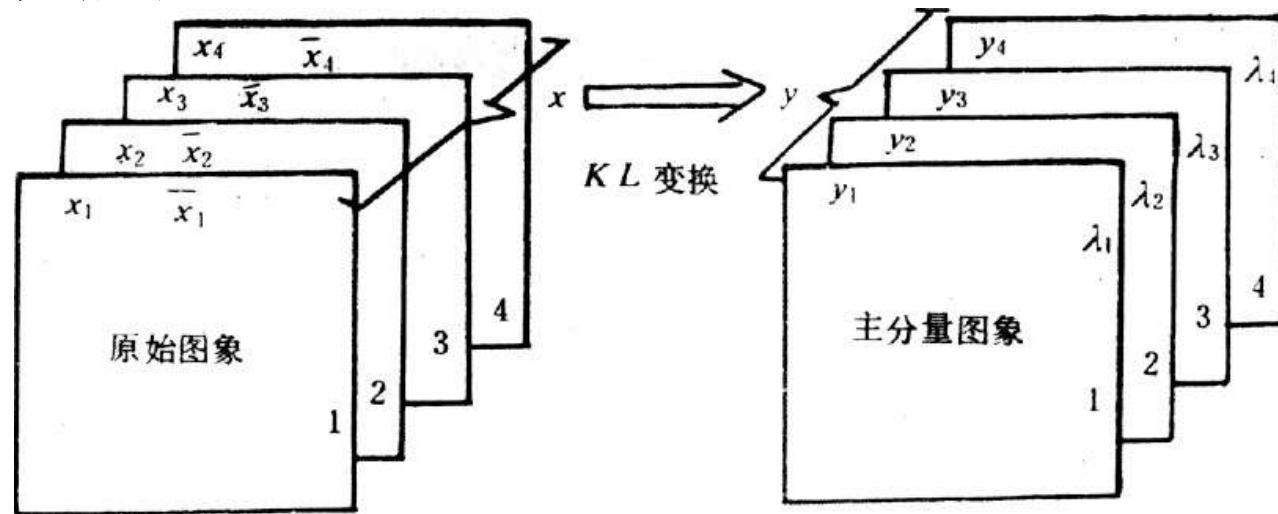
- 2.1 引言
- 2.2 空域变换
- 2.3 频率域变换
- 2.4 离散余弦变换
- 2.5 K-L 变换**
- 2.6 其他正交变换



第二章 常用的数学变换

2.5 K-L 变换

- 各主分量相互之间互不相关，使得变换后的特征具有独立描述图像的全部信息
- 在遥感图像处理中，将原始的多波段图像转换为主分量图像，将大量信息更多地集中在少数几个分量上，实现有效的特征抽取





第二章 常用的数学变换

2.5 K-L 变换

KL 变换（主成分分析）在遥感图像中的应用举例：美国南卡罗来纳州查尔斯顿地区 TM 影像的主成分分析—引于 John R. Jensen 著

《遥感数字影像处理导论》



b. Landsat TM bands 4, 3, 2 = RGB.



第二章 常用的数学变换

2.5 K-L 变换

□ 用于主成分分析的统计量

Band Number:	1	2	3	4	5	7	6
μm:	0.45-0.52	0.52-0.60	0.63-0.69	0.76-0.90	1.55-1.75	2.08-2.35	10.4-12.5
单变量统计							
Mean	64.80	25.60	23.70	27.30	32.40	15.00	110.60
Standard Deviation	18.05	5.84	8.30	15.76	23.85	12.45	4.21
Variance	100.93	34.14	68.83	248.40	568.84	154.92	17.78
Minimum	51	17	14	4	0	0	90
Maximum	242	115	131	105	193	128	130
协方差矩阵							
1	100.93						
2	56.68	34.14					
3	79.43	46.71	68.83				
4	61.49	40.38	69.59	248.40			
5	134.27	85.22	141.04	330.71	568.84		
7	90.13	55.14	86.91	142.50	280.97	154.92	
6	23.72	14.33	22.92	43.62	78.91	42.65	17.78





第二章 常用的数学变换

2.5 K-L 变换

□ 用于主成分分析的统计量

相关矩阵													
1	1.00												
2		0.96	1.00										
3			0.95	0.96	1.00								
4				0.39	0.44	0.53	1.00						
5					0.56	0.61	0.71	1.00					
7						0.73	0.76	0.84	0.76	0.95	1.00		
6							0.56	0.58	0.66	0.66	0.78	0.81	1.00



第二章 常用的数学变换

2.5 K-L 变换

□由协方差矩阵得到的特征值

	主成分						
	1	2	3	4	5	6	7
特征值	1010.92	131.20	37.60	6.73	3.95	2.17	1.24
Difference	879.72	93.59	30.88	2.77	1.77	.93	--
总方差	= 1193.81						

每个主成分解释的数据总方差的百分比由下式计算

$$\text{Computed as \%}_p = \frac{\text{eigenvalue } \lambda_p}{\sum \text{ eigenvalue } \lambda_p} \times 100$$

$$\sum \text{ eigenvalue } \lambda_p$$

例如：

$$\sum_{p=1}^7 \lambda_p = 1010.92 + 131.20 + 37.60 + 6.73 + 3.95 + 2.17 + 1.24 = 1193.81$$

$$\text{第一主成分解释的方差百分比} = \frac{1010.92 \times 100}{1193.81} = 84.68$$

百分比	84.68	10.92	3.15	0.56	0.33	0.18	0.10
累积百分比	84.68	95.67	98.82	99.38	99.71	99.89	99.99



第二章 常用的数学变换

2.5 K-L 变换

□由协方差矩阵得到的特征值

	主成分							
	1	2	3	4	5	6	7	
band 1	0.205	0.637	0.327	-0.054	0.349	-0.611	-0.879	
2	0.127	0.342	0.169	-0.077	0.012	0.396	0.821	
3	0.204	0.428	0.159	-0.076	-0.075	0.649	-0.562	
4	0.443	-0.471	0.739	0.107	-0.153	-0.019	-0.004	
5	0.742	-0.177	-0.427	-0.300	0.370	0.007	0.011	
7	0.376	0.197	-0.309	-0.312	-0.769	-0.181	0.051	
6	0.186	0.033	-0.020	0.827	0.434	0.122	0.005	





第二章 常用的数学变换



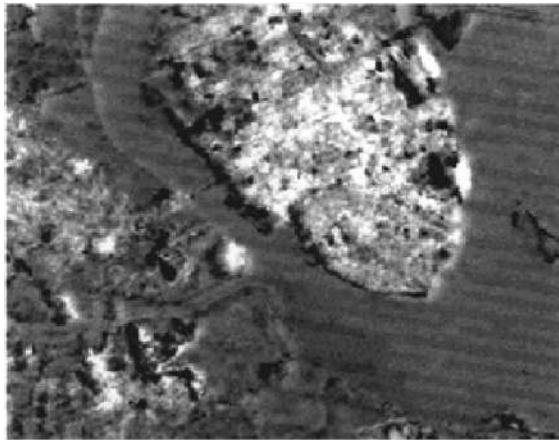
Principal component 1.



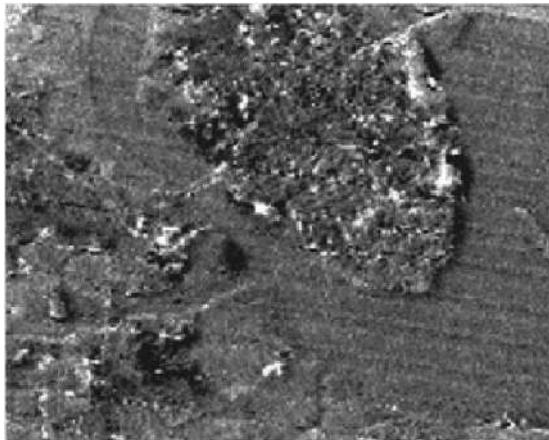
Principal component 2.



Principal component 3.



Principal component 4.



Principal component 5.



Principal component 6.





第二章 常用的数学变换

- 2.1 引言
- 2.2 空域变换
- 2.3 频率域变换
- 2.4 离散余弦变换
- 2.5 KL 变换
- 2.6 其他正交变换



第二章 常用的数学变换

2.6 其他正交变换

换

□ 能否进一步找到计算更简单的变换？

◆ 方法1：构建更为简单的正交变换集，只要满足正交关系：

$$[A][A]^T = I$$

◆ 对正弦函数集进行深入研究，发现不考虑函数值，仅考察函数值得过零点位置分布时，可形成包含+1和-1极值状态下的正交函数集。



第二章 常用的数学变换

2.6 其他正交变换 — 离散沃尔什 - 哈达玛变换 (WHT)

□ 沃尔什函数

沃尔什函数是 1923 年由美国数学家沃尔什提出的。沃尔什函数系是一完备正交函数系，其值只能取 +1 和 -1。从排列次序上可将沃尔什函数分为三种定义方法：一种是按照沃尔什排列来定义（按列率排序）；另一种是按照佩利排列来定义（按自然排序）；第三种是按照哈达玛排列来定义。由于哈达玛排序的沃尔什函数是由 2^n ($n=0, 1, 2, \dots$) 阶哈达玛矩阵（Hadamard Matrix）得到的，而哈达玛矩阵的最大优点在于它具有简单的递推关系，即高阶矩阵可用两个低阶矩阵的克罗内克积求得，因此在此只介绍哈达玛排列定义的沃尔什变换。



第二章 常用的数学变换

2.6 其他正交变换 — 离散沃尔什 - 哈达玛变换 (WHT)

- 1893 年法国数学家哈达玛总结前人研究只包含 +1 和 -1 的正交矩阵结果，形成哈达玛矩阵，既简单又有规律
- 1923 年美国数学家沃尔什提出 Walsh 函数，具有特点
 - ◆ 函数取值仅有两个 (0, 1 或 -1, +1)
 - ◆ 由 Walsh 函数构成的 Walsh 函数集，具备正交性和完备性
- 将 Walsh 函数用于信号的变换，建立 Walsh 变换基础
- 后人发现，按照哈达玛构造矩阵的排列方式，对 Walsh 函数进行排列，形成的韩书记既满足正交性和完备性，有特别容易记忆和产生，将该正交函数集应用于信号变换，由此形成常用的沃尔什 - 哈达玛变换



第二章 常用的数学变换

2.6 其他正交变换 — 离散沃尔什 - 哈达玛变换 (WHT)

□ $N=2^n (n=0, 1, 2, \dots)$ 阶哈达玛矩阵每一行的符号变化规律对应于某一个沃尔什函数的符号变化规律，即 $N=2^n (n=0, 1, 2, \dots)$ 阶哈达玛矩阵的每一行对应于一个离散沃尔什函数，哈达玛矩阵与沃尔什函数系不同之处仅仅是行的次序不同。 2^n 阶哈达玛变换矩阵 H 具有如下递推公式：

$$H_1 = [1], \quad H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad H_{2N} = \frac{1}{\sqrt{2}} \begin{bmatrix} H_N & H_N \\ H_N & -H_N \end{bmatrix}$$



第二章 常用的数学变换

2.6 其他正交变换 — 离散沃尔什 - 哈达玛变换 (WHT)

$$H_8 = \frac{1}{2\sqrt{2}}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & -1 & -1 \end{bmatrix}$$



第二章 常用的数学变换

2.6 其他正交变换 — 离散沃尔什 - 哈达玛变换 (WHT)

□ 一维离散沃尔什变换定义为

$$W(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) Walsh(u, x)$$

一维离散沃尔什逆变换定义为

$$f(x) = \sum_{u=0}^{N-1} W(u) Walsh(u, x)$$

式中， $Walsh(u, x)$ 为沃尔什函数。若将 $Walsh(u, x)$ 用哈达玛矩阵表示，并将变换表达式写成矩阵形式，则上式分别为



第二章 常用的数学变换

2.6 其他正交变换 — 离散沃尔什 - 哈达玛变换 (WHT)

$$\begin{bmatrix} W(0) \\ W(1) \\ \vdots \\ W(N-1) \end{bmatrix} = \frac{1}{N} [H_N] \begin{bmatrix} f(0) \\ f(1) \\ \vdots \\ f(N-1) \end{bmatrix} \text{ 和 } \begin{bmatrix} f(0) \\ f(1) \\ \vdots \\ f(N-1) \end{bmatrix} = [H_N] \begin{bmatrix} W(0) \\ W(1) \\ \vdots \\ W(N-1) \end{bmatrix}$$

式中, $[H_N]$ 为 N 阶哈达玛矩阵

由哈达玛矩阵的特点可知, 沃尔什 - 哈达玛变换的本质上是将离散序列 $f(x)$ 的各项值的符号按一定规律改变后, 进行加减运算, 因此, 它比采用复数运算的 DFT 和采用余弦运算的 DCT 要简单得多。



第二章 常用的数学变换

2.6 其他正交变

换二维离散沃尔什 - 哈达玛变换◆

很容易将一维 WHT 的定义推广到二维 WHT。二维 WHT 的正变换

核和逆变换核分别为：

$$W(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) Walsh(u, x) Walsh(v, y)$$

和

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} W(u, v) Walsh(u, x) Walsh(v, y)$$

式中： $x, u=0, 1, 2, \dots, M-1 ; y, v=0, 1, 2, \dots, N-1$ 。



第二章 常用的数学变换

2.6 其他正交变换

$$f_1 = \begin{bmatrix} 1 & 3 & 3 & 1 \\ 1 & 3 & 3 & 1 \\ 1 & 3 & 3 & 1 \\ 1 & 3 & 3 & 1 \end{bmatrix}$$

和

$$f_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

求这两个图像的二维 WHT。

根据题意，知 $M=N=4$ ，其二维 WHT 变换核为

$$H_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$



第二章 常用的数学变换

2.6 其他正交变换

$$W_1 = \frac{1}{4^2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 & 1 \\ 1 & 3 & 3 & 1 \\ 1 & 3 & 3 & 1 \\ 1 & 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$



第二章 常用的数学变换

2.6 其他正交变换

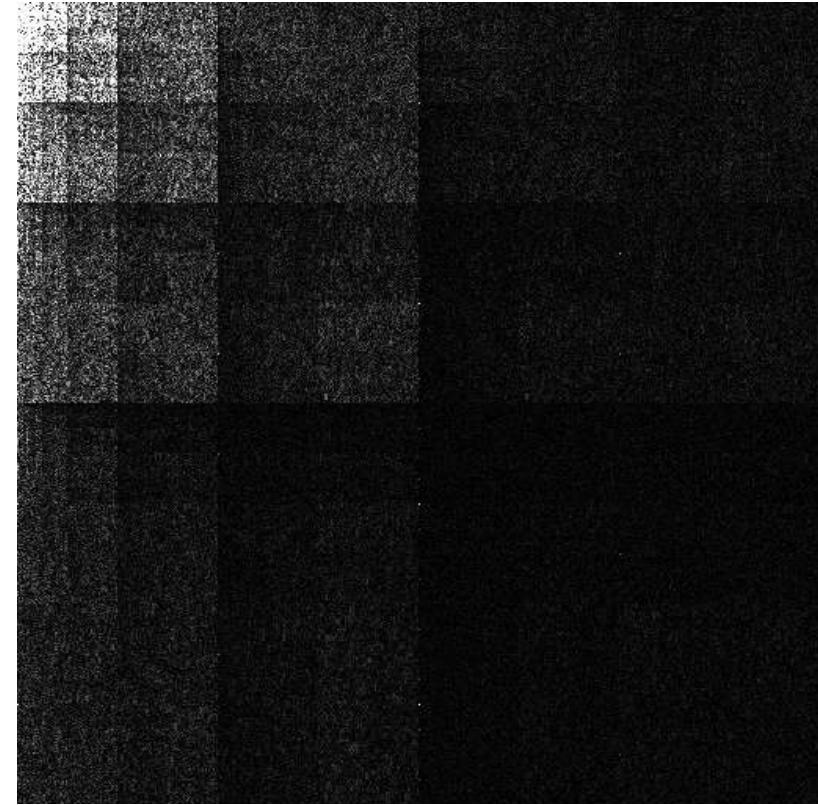
$$W_2 = \frac{1}{4^2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$





第二章 常用的数学变换

2.6 其他正交变换





第二章 常用的数学变换

- 2.1 引言
- 2.2 空域变换
- 2.3 频率域变换
- 2.4 离散余弦变换
- 2.5 KL 变换
- 2.6 其他正交变换



第二章 常用的数学变换

第二章结
束谢 谢！



数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第3章 图像增强

第1讲 图像增强的概念、灰度直方图及线性变换

演示文稿说明：

- 本讲内容以板书为主，*ppt* 演示为辅；
- 本讲部分图片来自冈萨雷斯的数字图像处理（英文版）教材。



第3章 图像增强

3.1 引言

3.2 空域增强

3.3 频域增强

3.4 彩色增强



第3章 图像增强

3.1 引言

图像增强 (Image enhancement) 是图像处理的基本内容之一。图像增强是指按特定的需要突出一幅图像中的某些信息，同时，削弱 (weaken) 或去除 (eliminate) 某些不需要的信息。这类处理是为了某种应用目的去改善图像质量，处理的结果更适合于人的视觉特性或机器识别系统。增强处理并不能增加原始图像的信息，而只能增强对某种信息的辨识能力，并且这种处理有可能损失一些其它信息。



第3章 图像增强

3.1 引言

图像增强处理技术基本上可以分成两大类，一类是空域处理方法，一类是频域处理方法。空域法是直接对图像中的像素进行处理，基本上是以灰度映射 (mapping) 变换为基础的，所用的映射变换取决于增强的目的。例如增加图像的对比度 (contrast)，改善图像的灰度层次等处理均属空域法处理。频域处理方法的基础是卷积定理 (convolution theorem)，它采用修改图像傅立叶变换的方法实现对图像的增强处理。



第3章 图像增强

3.1 引言

空域增强

点运算

邻域运算

几何变换

彩色增强

线性

非线性：中值滤波等

图像平滑

图像锐化

频域增强（滤波）：高通、低通、带通、带阻等



第3章 图像增强

3.2 空域增强

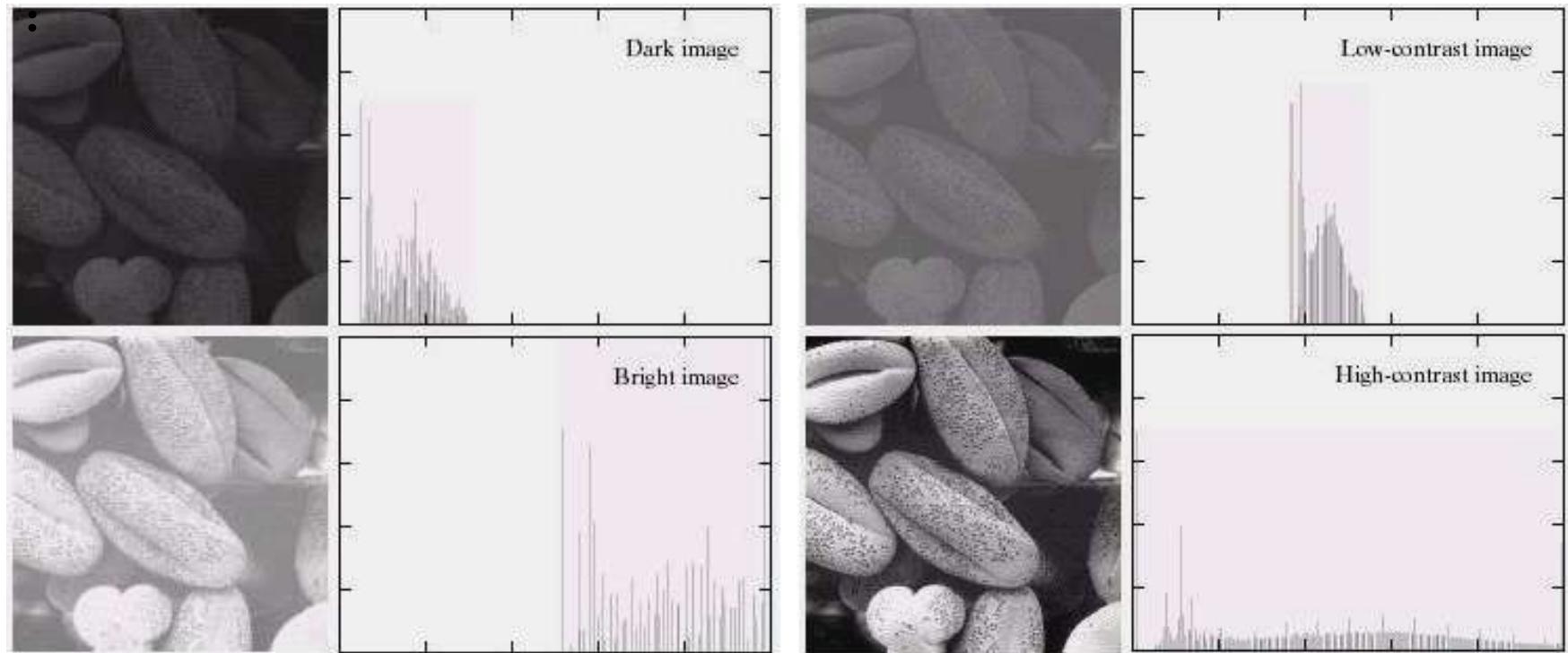
灰度直方图：对一幅数字图像，若对应于每一灰度值，统计出具有该灰度值的象素数，并据此绘出象素数-灰度值图形，则该图形称该图像的灰度直方图，简称直方图。直方图是以灰度值作横坐标，象素数作纵坐标。有时直方图亦采用某一灰度值的象素数占全图总象素数的百分比（即某一灰度值出现的频数）作为纵坐标。



第3章 图像增强

3.2 空域增强

不同亮度图像的灰度直方图



图片来自冈萨雷斯的数字图像处理(英文版)教材



第3章 图像增强

3.2 空域增强

直方图的性质：

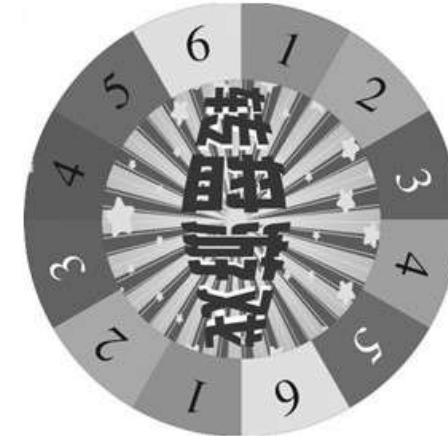
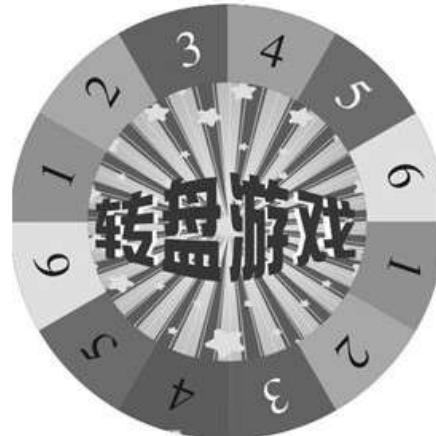
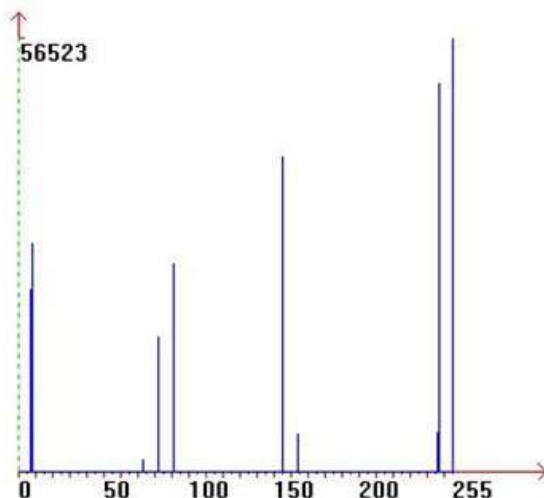
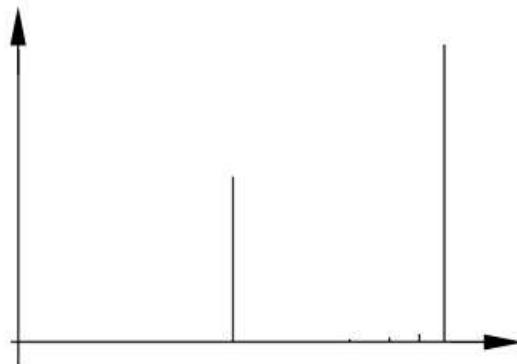
- 直方图是一幅图像中各象素灰度值出现次数（或频数）的统计结果，它只反映该图像中不同灰度值出现的次数（或频数），而未反映每一灰度值象素所在位置。也就是说，它只包含了该图像中某一灰度值的象素出现的概率，而丢失了其所在位置的信息。
- 任一幅图像，都能唯一地算出一幅与它对应的直方图，但不同的图像，可能有相同的直方图。也就是说，图像与直方图之间是一种多对一的映射关系。
- 由于直方图是对具有相同灰度值的象素统计计数得到的，因此，一幅图像各子区的直方图之和就等于该图全图的直方图。



第3章 图像增强

3.2 空域增强

直方图的性质：





第3章 图像增强

3.2 空域增强

灰度变换：

设T操作为灰度变换函数(强度变换)，则有：

$$s = T(r)$$

s 为 $g(x,y)$ 象素的灰度值； r 为 $f(x,y)$ 象素的灰度值。

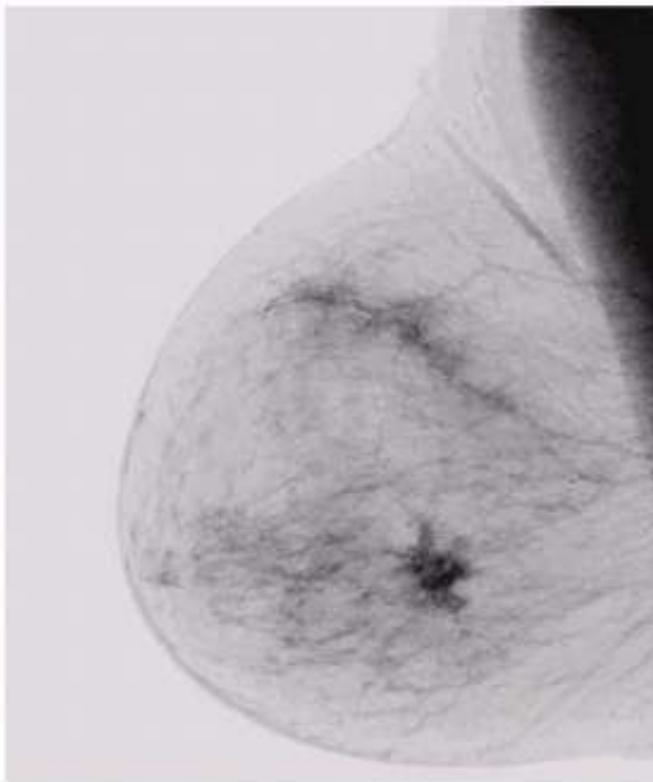
灰度变换分为线性变换和非线性变换。



第3章 图像增强

3.2 空域增强

灰度线性变换(反色):



a b

FIGURE 3.4
(a) Original digital mammogram.
(b) Negative image obtained using the negative transformation in Eq. (3.2-1).
(Courtesy of G.E. Medical Systems.)

图片来自冈萨雷斯的数字图像处理(英文版)教材



第3章 图像增强

3.2 空域增强

灰度线性变换(反色):



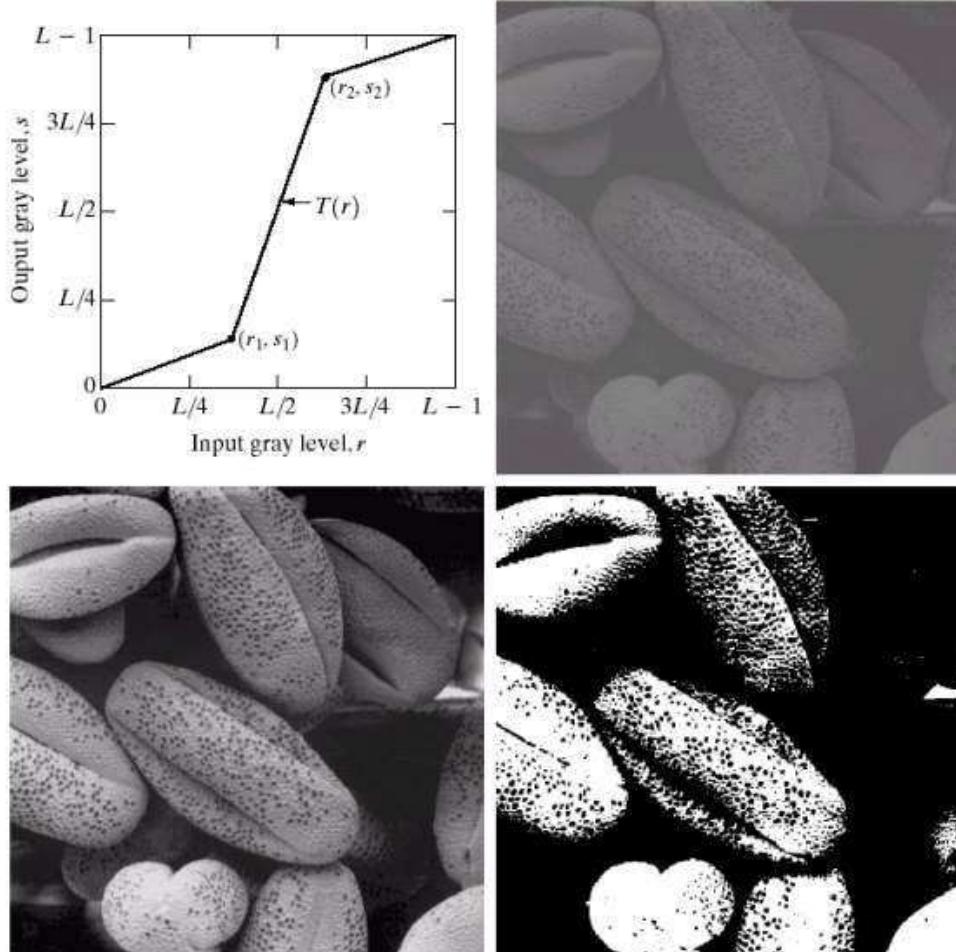
数字图像处理



第3章 图像增强

3.2 空域增强

灰度线性变换(对比度扩展):



a
b
c
d

FIGURE 3.10
Contrast stretching.
(a) Form of transformation function.
(b) A low-contrast image.
(c) Result of contrast stretching.
(d) Result of thresholding.
(Original image courtesy of Dr. Roger Heady, Research School of Biological Sciences, Australian National University, Canberra, Australia.)

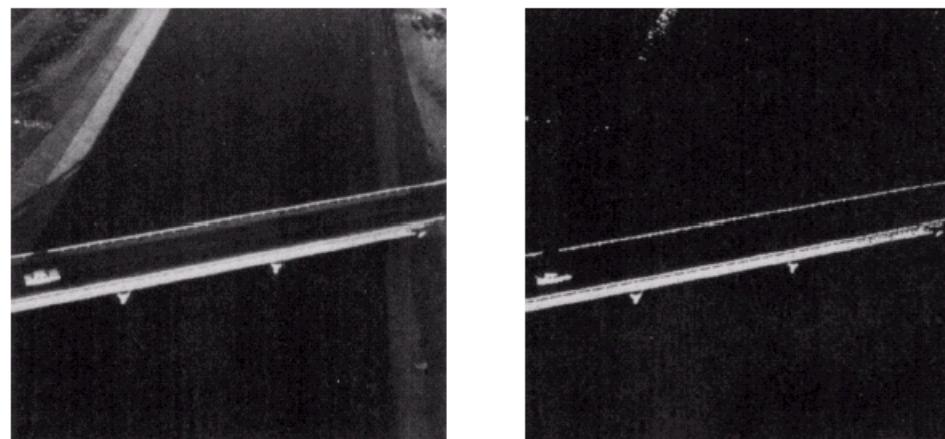
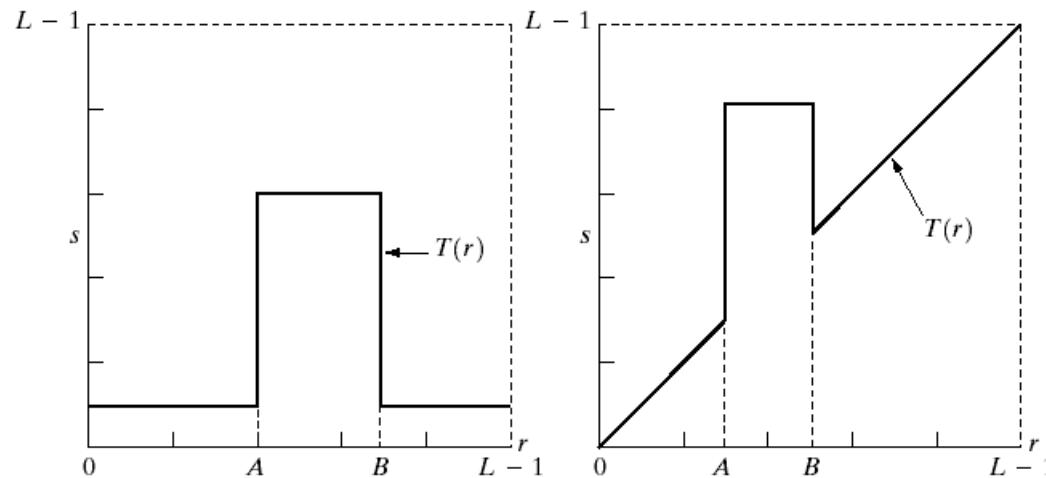
图片来自冈萨雷斯的数字图像处理(英文版)教材



第3章 图像增强

3.2 空域增强

灰度线性变换(灰度切分):



a b
c d

FIGURE 3.11
(a) This transformation highlights range $[A, B]$ of gray levels and reduces all others to a constant level.
(b) This transformation highlights range $[A, B]$ but preserves all other levels.
(c) An image.
(d) Result of using the transformation in (a).

图片来自冈萨雷斯的数字图像处理(英文版)教材





第3章 图像增强

3.2 空域增强

位图切割:

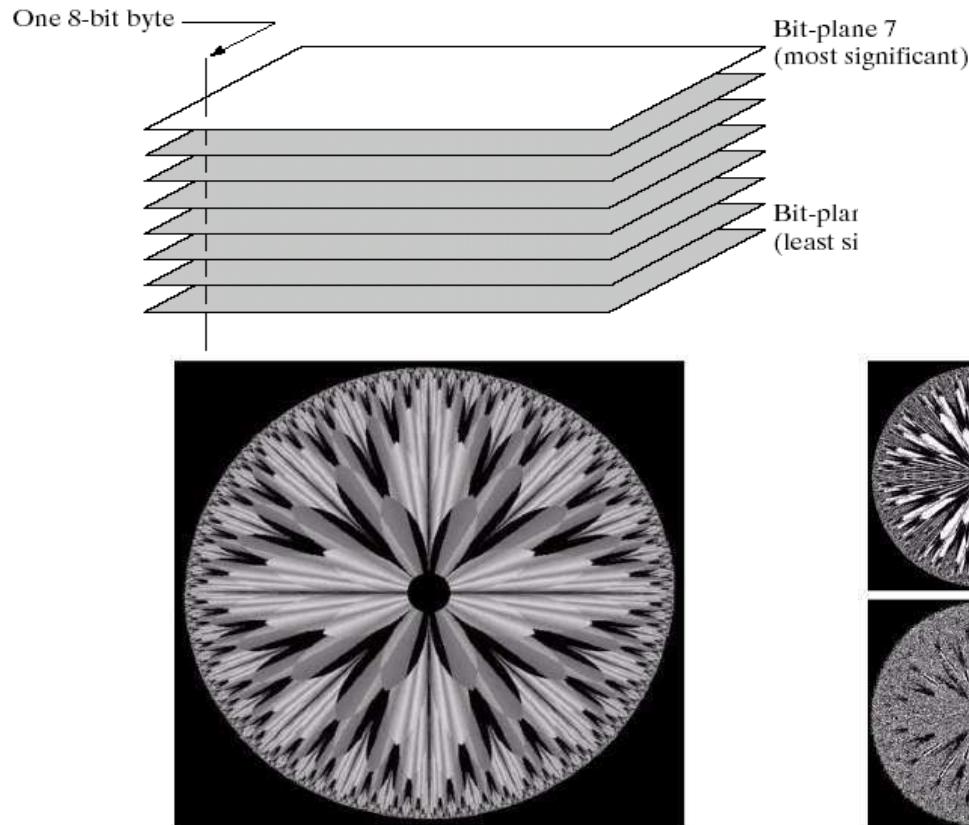


FIGURE 3.13 An 8-bit fractal image. (A fractal is an image generated from mathematical expressions). (Courtesy of Ms. Melissa D. Binde, Swarthmore College, Swarthmore, PA.)

FIGURE 3.12
Bit-plane
representation of
an 8-bit image.

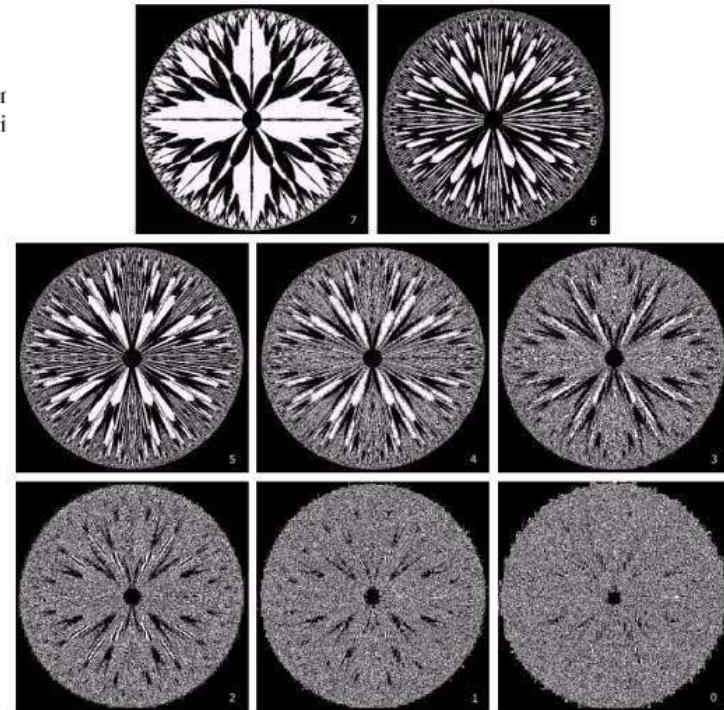


FIGURE 3.14 The eight bit planes of the image in Fig. 3.13. The number at the bottom, right of each image identifies the bit plane.

图片来自冈萨雷斯的数字图像处理(英文版)教材



数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第3章 图像增强

第2讲 灰度非线性变换及直方图均衡化

演示文稿说明：

- 本讲内容以板书为主，*ppt* 演示为辅；
- 本讲部分图片来自冈萨雷斯的数字图像处理（英文版）教材。

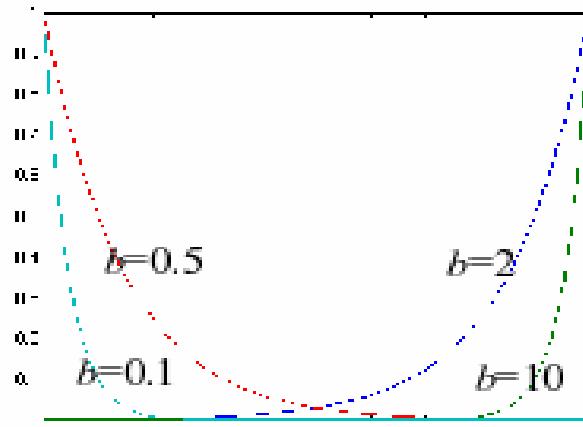


第3章 图像增强

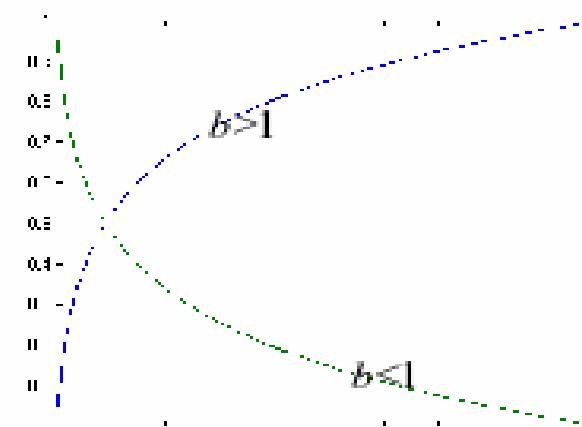
3.2 空域增强

灰度非线性变换：

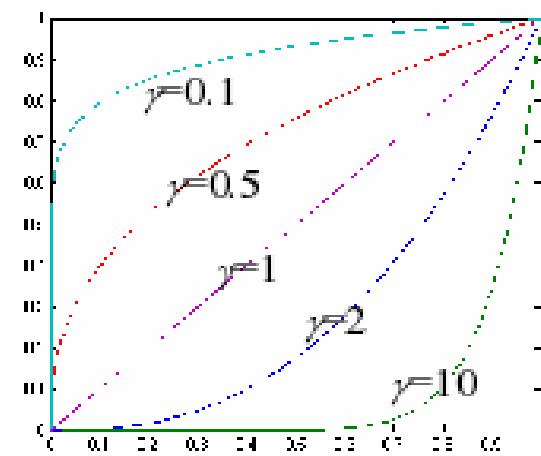
- 指数函数变换
- 对数函数变换
- 幂函数变换



(a) 指数函数



(b) 对数函数



(c) 幂函数



第3章 图像增强

3.2 空域增强

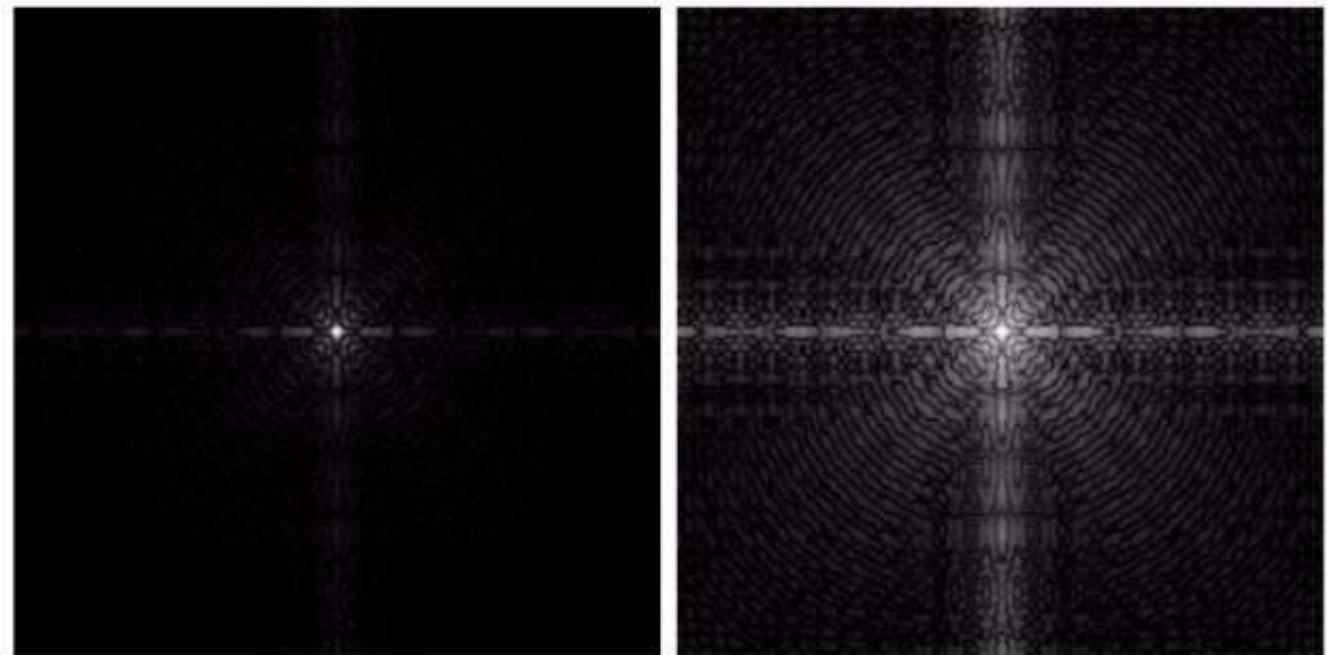
灰度非线性变换

对数变换示例: $s = c \log(1 + r)$

a b

FIGURE 3.5

(a) Fourier spectrum.
(b) Result of applying the log transformation given in Eq. (3.2-2) with $c = 1$.



图片来自冈萨雷斯的数字图像处理（英文版）教材。



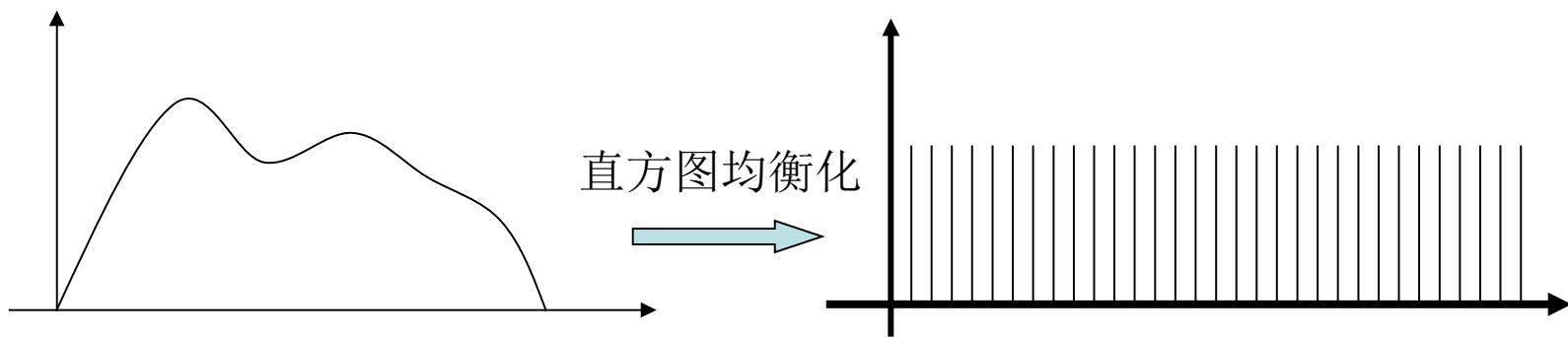


第3章 图像增强

3.2 空域增强

直方图均衡化：是将原图像通过某种变换，得到一幅灰度直方图为均匀分布的新图像的方法。

结果：图像的对比度得到增强。





第3章 图像增强

3.2 空域增强

设 r 和 s 分别表示归一化了的原图像灰度和经直方图均衡化后的图像灰度，即

$$0 \leq r, s \leq 1$$

在 $[0,1]$ 区间内的任一个 r 值，都可产生一个 s 值，且

$$s = T(r)$$

$T(r)$ 作为变换函数，满足下列条件：

- ① 在 $0 \leq r \leq 1$ 内为单调递增函数，保证灰度级从黑到白的次序不变；
- ② 在 $0 \leq r \leq 1$ 内，有 $0 \leq T(r) \leq 1$ ，确保映射后的像素灰度在允许的范围内。



第3章 图像增强

3.2 空域增强

反变换关系为

$$r = T^{-1}(s)$$

$T^{-1}(s)$ 对 s 同样满足上述两个条件。

由概率论理论可知，如果已知随机变量 r 的概率密度为 $p_r(r)$ ，而随机变量 s 是 r 的函数，则 s 的概率密度 $p_s(s)$ 可以由 $p_r(r)$ 求出。

假定随机变量 s 的分布函数用 $F_s(s)$ 表示，根据分布函数定义 $F_s(s) = \int_{-\infty}^s p_s(s) ds = \int_{-\infty}^r p_r(r) dr$



第3章 图像增强

3.2 空域增强

利用密度函数是分布函数的导数关系，等式两边对 s 求导：

$$P_s(s) = \frac{d}{ds} \left[\int_{-\infty}^r p_r(r) dr \right] = p_r \frac{dr}{ds} = p_r \frac{d}{ds} [T^{-1}(s)]$$

可见，输出图像的概率密度函数可以通过变换函数 $T(r)$ 控制原图像灰度级的概率密度函数得到，因而改善原图像的灰度层次，这就是直方图均衡化技术的基础。

从人眼视觉特性来考虑，一幅图像的直方图如果是均匀分布的，即 $P_s(s)=k$ （归一化时 $k=1$ ）时，该图像色调给人的感觉比较协调。因此将原图像直方图通过 $T(r)$ 调整为均匀分布的直方图，这样修正后的图像能满足人眼视觉要求。

归一化假定 $ds = p_r(r)dr$

则有





第3章 图像增强

3.2 空域增强

两边积分得

$$s = T(r) = \int_0^r p_r(r) dr$$

上式表明，当变换函数为 r 的累积直方图函数时，能达到直方图均衡化的目的。

对于离散的数字图像，设总像素数为 N，灰度级总数为 L，第 K 个灰度级的值为 r_k ，图像中具有灰度级 r_k 的像素数为 n_k ，则变换函数 $T(r_k)$ 的离散形式可表示为：

$$s_k = T(r_k) = \sum_{j=0}^k p_r(r_j) = \sum_{j=0}^k \frac{n_j}{N}$$

上式表明，均衡后各像素的灰度值 s_k 可直接由原图像的直方图算出。



第3章 图像增强

3.2 空域增强

例题 3.1 假设一幅大小为 64*64 的图像，8 个灰度级，灰度分布如下表，将其灰度均衡化。

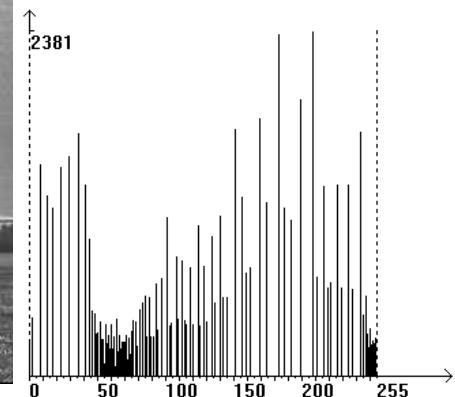
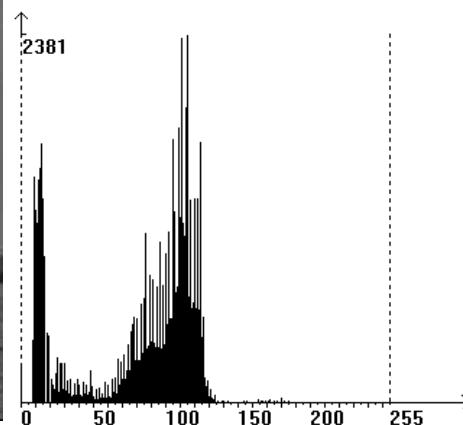


第3章 图像增强

3.2 空域增强

灰度直方图增强

直方图均衡化：





数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第3章 图像增强

第3讲 直方图规定化

演示文稿说明：

- 本讲内容以板书为主，*ppt* 演示为辅；



第3章 图像增强

3.2 空域增强

直方图规定化：

变换直方图使之成为所要求的形状，从而有选择地增强某个灰度值范围。

$P_r(r_k)$: 原图的灰度概率密度函数；

$P_z(z_k)$: 规定化后的灰度概率密度函数。

对 r 、 z 分别进行均衡化处理：

$$s = T(r) = \int_0^r p_r(l) dl \quad u = G(z) = \int_0^z p_z(l) dl$$

s 与 u 相对应。



第3章 图像增强

3.2 空域增强

$$r = T^{-1}(s)$$

$$z = G^{-1}(u) = G^{-1}(s) = G^{-1}(T(r))$$

从而得到 z 与 r 间的对应关系。

对于离散图像: $p_r(r_k) = n_k / N$

$$u_k = G(z_k) = \sum_{j=0}^k p_z(z_j)$$

$$s_k = T(r_k) = \sum_{j=0}^k p_r(r_j) = \sum_{j=0}^k \frac{n_j}{N}$$

$$z_k = G^{-1}(u_k) = G^{-1}(s_k) = G^{-1}(T(r_k))$$



第3章 图像增强

3.2 空域增强

例题 3.2 设一幅图像具有如表 3.1 所示的概率分布，要求对其进行直方图规定化，规定化后的图像具有如表 3.2 所示的灰度级分布。

表 3.1

灰度级	0	1	2	3	4	5	6	7
对应概率分布	0.19	0.25	0.21	0.16	0.08	0.06	0.03	0.02

表 3.2

灰度级	0	1	2	3	4	5	6	7
对应概率分布	0	0	0	0.15	0.20	0.30	0.20	0.15



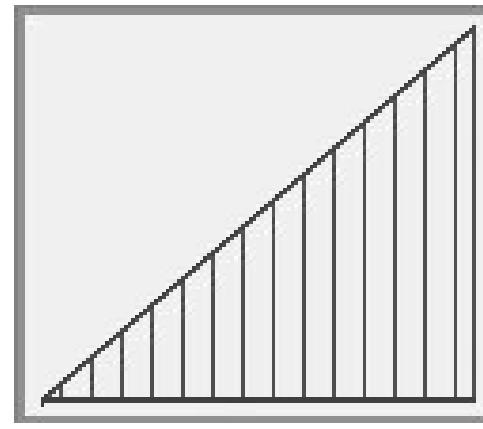
第3章 图像增强

3.2 空域增强

直方图规定化：



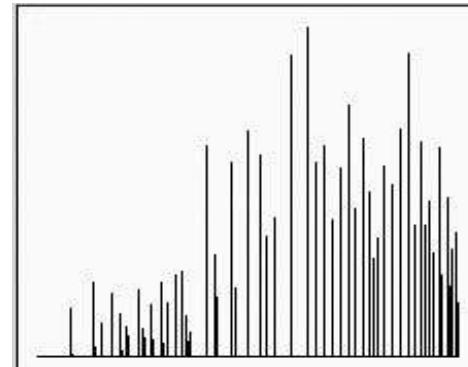
原图



规范化函数



直方图规定化



直方图



第3章 图像增强

3.2 空域增强

- ✓ 邻域运算

邻域平均法：

$$g(x, y) = [f(x, y) + \sum_{(i, j) \in R} f(i, j)] / (m + 1)$$

模板的卷积运算实现：



数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第3章 图像增强

第4讲 邻域平均法及中值滤波

演示文稿说明：

- 本讲内容以板书为主，*ppt* 演示为辅；
- 本讲部分图片来自冈萨雷斯的数字图像处理（英文版）教材。



第3章 图像增强

3.2 空域增强

邻域平均法：是一种直接在空间域上进行平滑处理的技术。假设图像是由许多灰度恒定的小块组成，相邻像素间存在很高的空间相关性，而噪声则是统计独立的。因此，可用邻域内各像素的灰度平均值代替该像素原来的灰度值，实现图像的平滑。

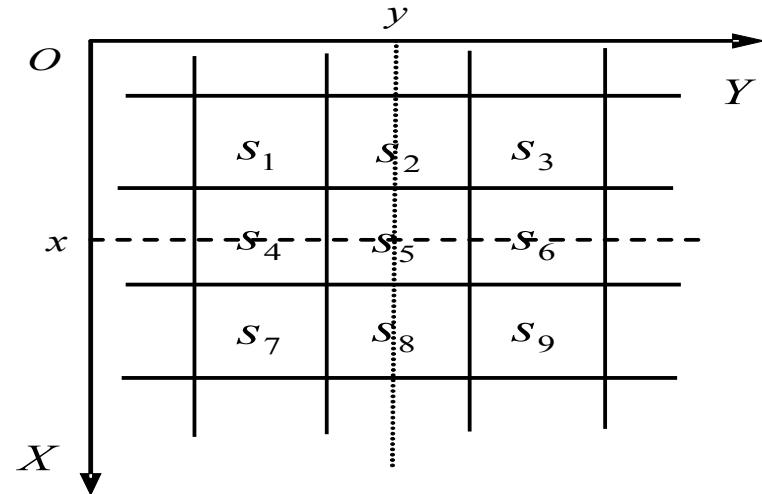


第3章 图像增强

3.2 空域增强

$$g(x, y) = w_1 s_1 + w_2 s_2 + \cdots + w_9 s_9$$

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9





第3章 图像增强

3.2 空域增强

邻域平均: $H = \frac{1}{4} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ $H = \frac{1}{8} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

邻域加权平均: $H = \frac{1}{5} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ $H = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

$H = \frac{1}{6} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ $H = \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ $H = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$

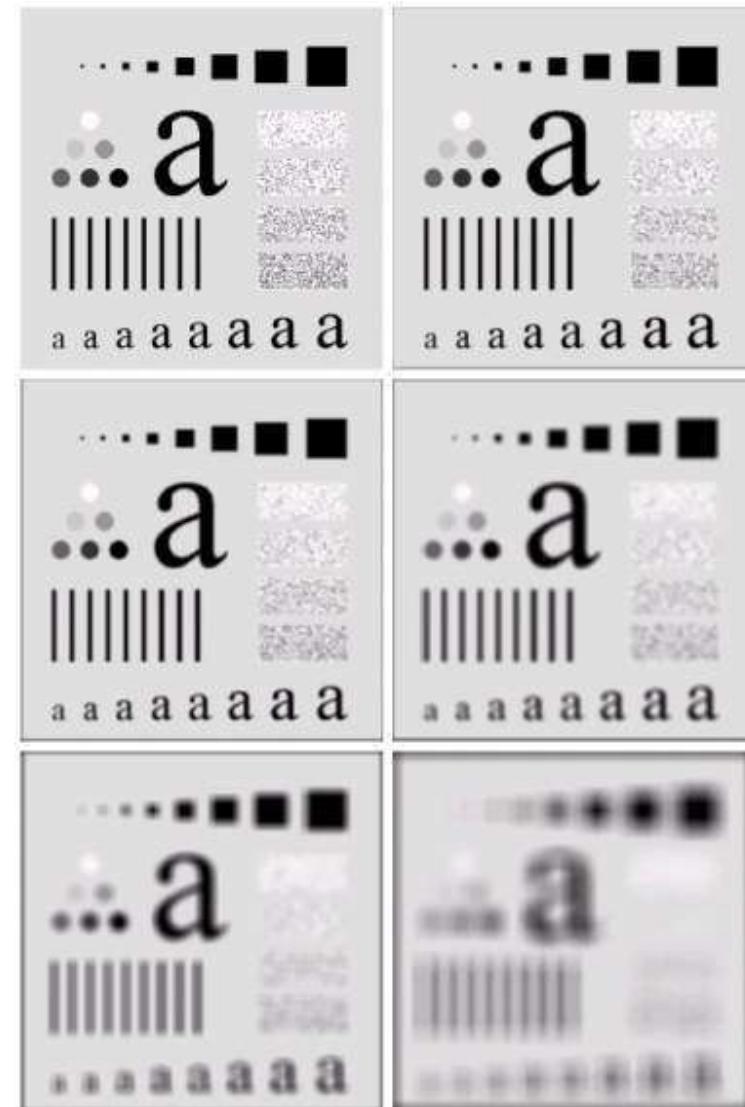


第3章 图像增强

3.2 空域增强

模板尺寸对图像平滑的效果

- 不同的模板尺寸：
 3×3 , 5×5 , 7×7 , ...
如右图。
- 有何影响？





第3章 图像增强

3.2 空域增强

中值滤波：将模板区域内的象素排序，求出中值。

中值滤波器是一种非线性平滑滤波器，既可消除噪声又可保持图象的细节。它的工作步骤如下：

- 将模板在图中漫游，并将模板中心与图中某个象素位置重合
- 读取模板下各对应象素的灰度值；
- 将这些灰度值从小到大排成**1**列；
- 找出这些值里排在中间的**1**个；
- 将这个中间值赋给对应模板中心位置的象素。



第3章 图像增强

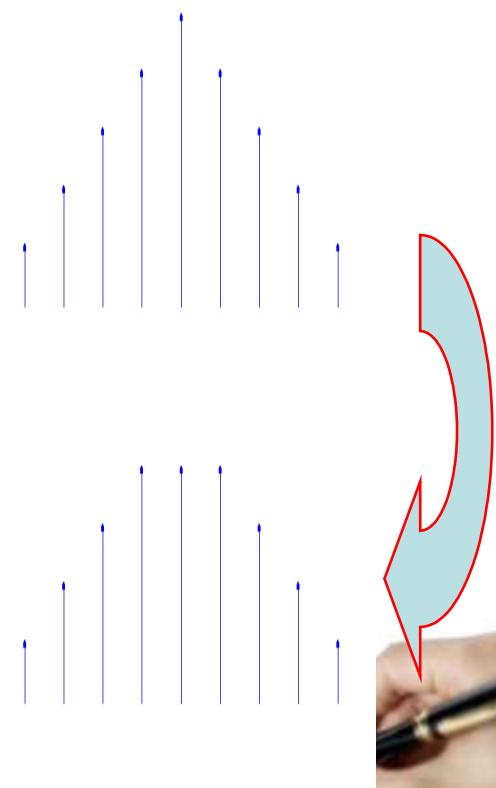
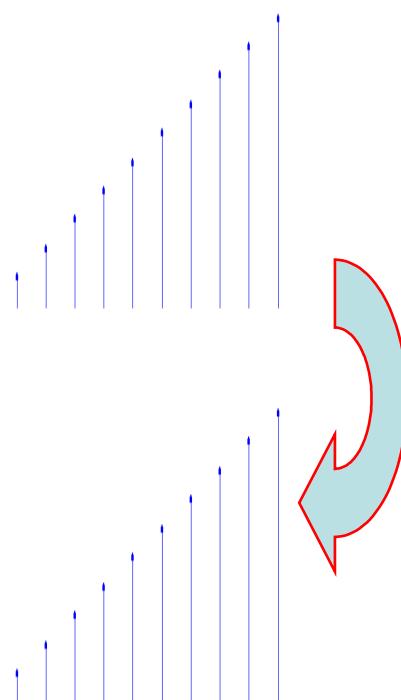
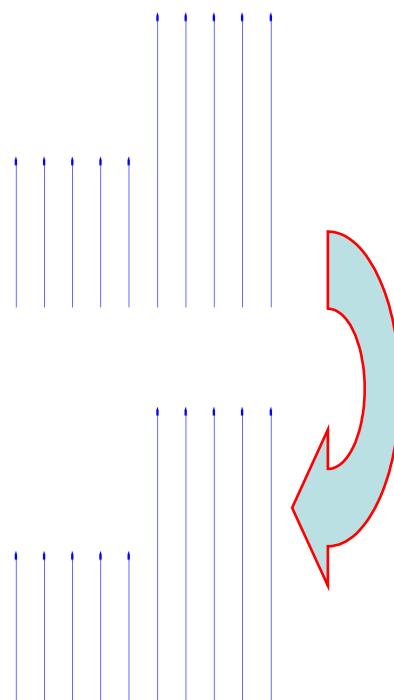
3.2 空域增强

中值滤波对小于窗宽 ($N=5$) $1/2$ 的脉冲将进行抑制，如：

阶跃

斜坡

三角

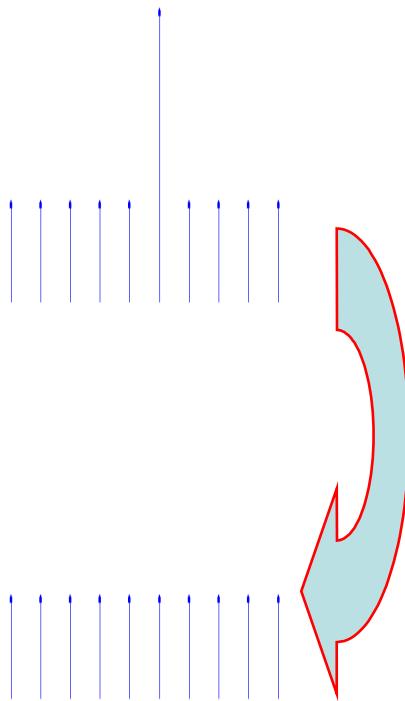




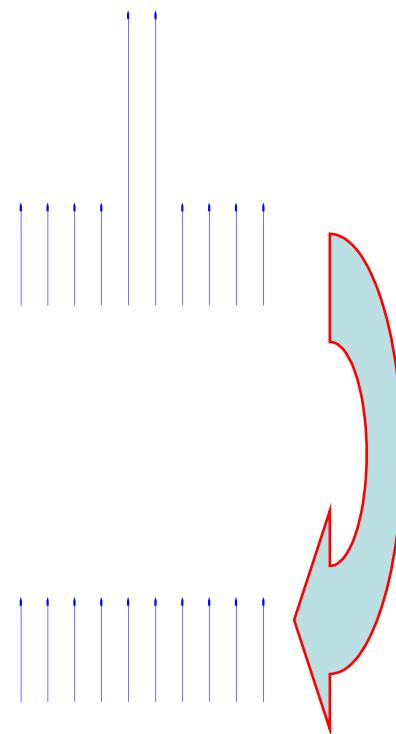
第3章 图像增强

3.2 空域增强

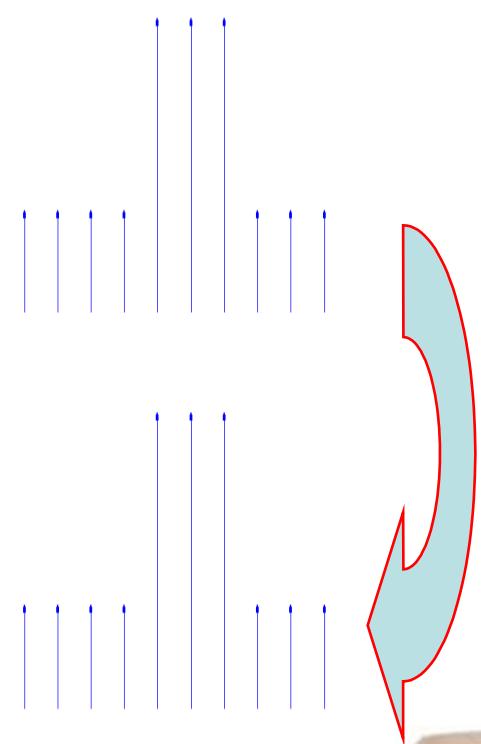
单脉冲



双脉冲



三脉冲

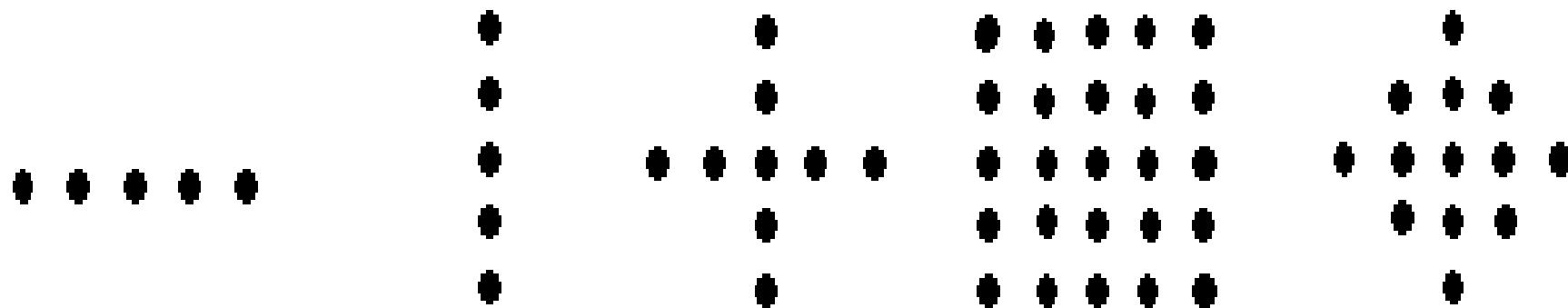




第3章 图像增强

3.2 空域增强

二维中值滤波器比一维滤波器有更好的性能。二维中值滤波器的窗口是二维的，其形状有很多种，如直线形、方形、圆形、十字形、圆环形、菱形等



(a) 水平直线形

(b) 垂直线形

(c) 十字形

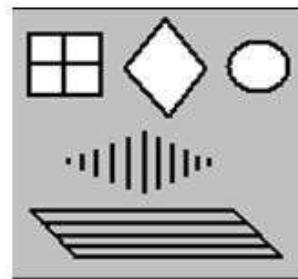
(d) 方形

(e) 菱形

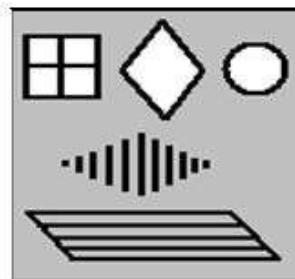


第3章 图像增强

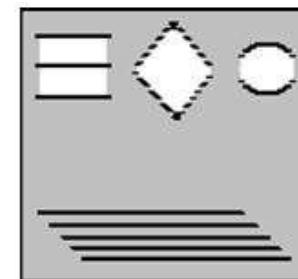
3.2 空域增强



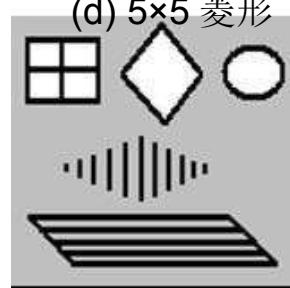
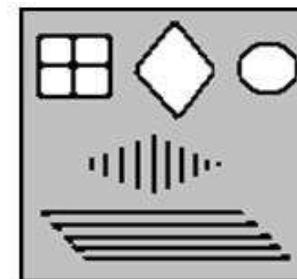
(a) 原图
(d) 5×5 菱形



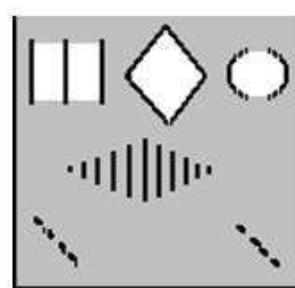
(b) 1×3 水平线形



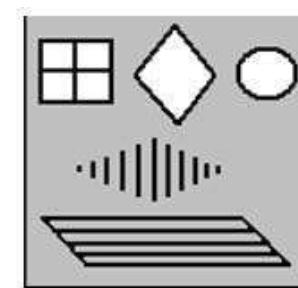
(c) 1×5 水平线形



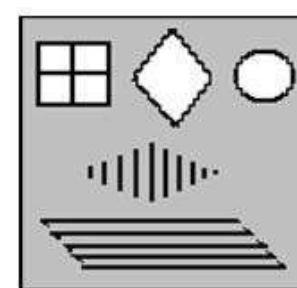
(e) 3×1 垂直线性
形



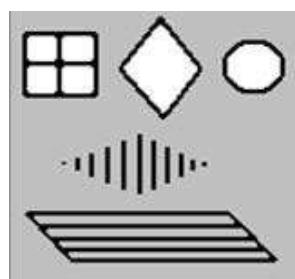
(f) 5×1 垂直线形



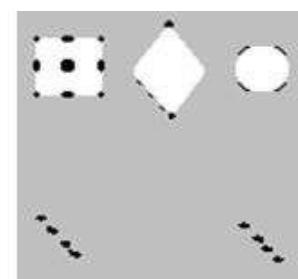
(g) 3×3 十字形



(h) 5×5 十字



(i) 3×3 方形



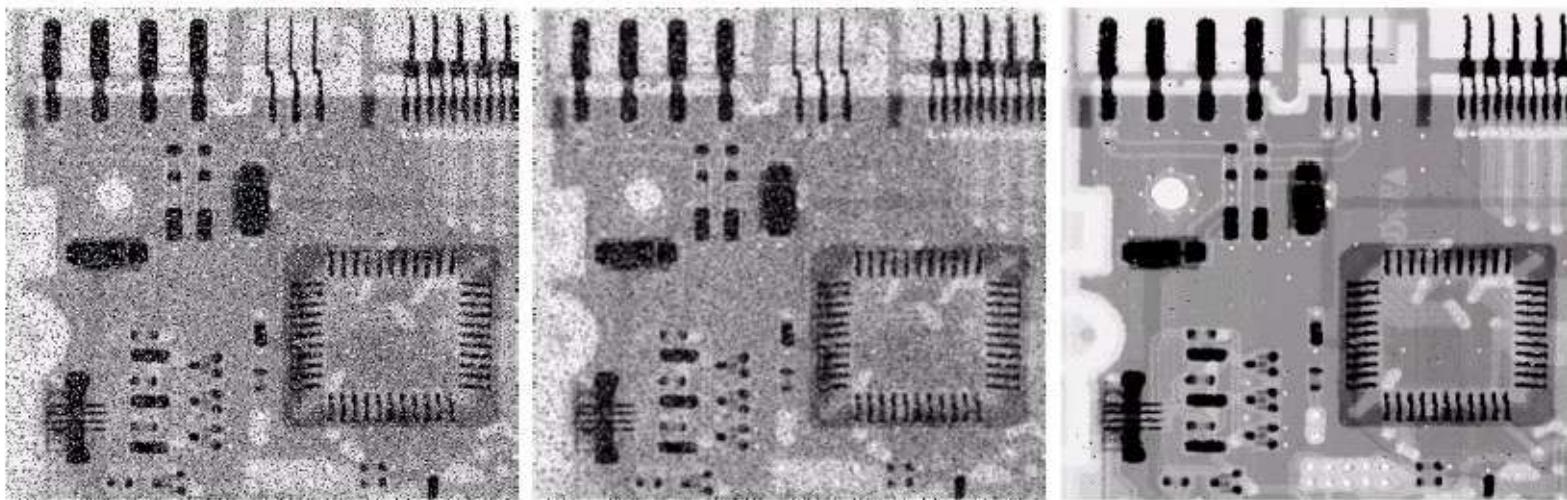
(j) 5×5 方形





3.2 空域增强

平均模版与中值滤波效果：



a b c

FIGURE 3.37 (a) X-ray image of circuit board corrupted by salt-and-pepper noise. (b) Noise reduction with a 3×3 averaging mask. (c) Noise reduction with a 3×3 median filter. (Original image courtesy of Mr. Joseph E. Pascente, Lixi, Inc.)



第3章 图像增强

3.2 空域增强

百分比滤波器：是一类更广泛的非线性滤波器，在工作的时候与中值滤波器相似。首先对窗口内的像素按灰度值的大小进行排序，然后根据某个确定的百分比选取序列中对应的像素值赋给窗口中心位置的像素。

如果百分比取最大就是**最大值滤波器**，它可以用来检测图像中的最亮点；

如果百分比取最小，就是**最小值滤波器**，它可以用来检测图像中的最暗；

如果百分比取 50%，就是**中值滤波器**。



第3章 图像增强

3.2 空域增强

空域锐化：

平滑 —— 积分运算

锐化 —— 微分运算

一阶偏导：

$$f_x(x, y) = f(x, y) - f(x-1, y)$$

$$f_y(x, y) = f(x, y) - f(x, y-1)$$



数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第3章 图像增强

第5讲 微分算子及拉普拉斯锐化

演示文稿说明：

- 本讲内容以板书为主，*ppt* 演示为辅；
- 本讲部分图片来自冈萨雷斯的数字图像处理（英文版）教材。



第3章 图像增强

3.2 空域增强

一阶微分算子

$f(x, y)$ 在点 (x, y) 处的梯度矢量为: $\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix}$

模:

$$|\nabla f_{(2)}| = \text{mag}(\nabla f) = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$$

简化为:

$$\left| \frac{\partial f}{\partial x} \right| + \left| \frac{\partial f}{\partial y} \right|$$



第3章 图像增强

3.2 空域增强

一阶微分算子

Z_1	Z_2	Z_3
Z_4	Z_5	Z_6
Z_7	Z_8	Z_9

$$G_x = Z_8 - Z_5$$

$$G_y = Z_6 - Z_5$$



第3章 图像增强

3.2 空域增强

Roberts 交叉微分算子

1	0
0	-1

0	1
-1	0

3x3 的 Prewitt 算子

-1	-1	-1
0	0	0
1	1	1

-1	0	1
-1	0	1
-1	0	1

3x3 的 Sobel 算子

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1



第3章 图像增强

3.2 空域增强

Laplacian 算子是不依赖于边缘方向的二阶微分算子，是常用的二阶导数算子。对一个连续函数 $f(x,y)$ ，它在位置 (x,y) 的 Laplacian 表示式：

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$



第3章 图像增强

3.2 空域增强

二阶微分（拉氏算子）

0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
-1	-1	-1

3x3 的拉氏算子



第3章 图像增强

3.2 空域增强

如何处理负值？

微分（差分）运算，以及剪影运算，结果都会出现负值，如何处理？

- (1) 将负值改为 0
- (2) 取绝对值（适合于提取边缘）
- (3) 加 255，再除以 2
- (4) 加最小值的绝对值，再将灰度值范围映射到 0-255 之间



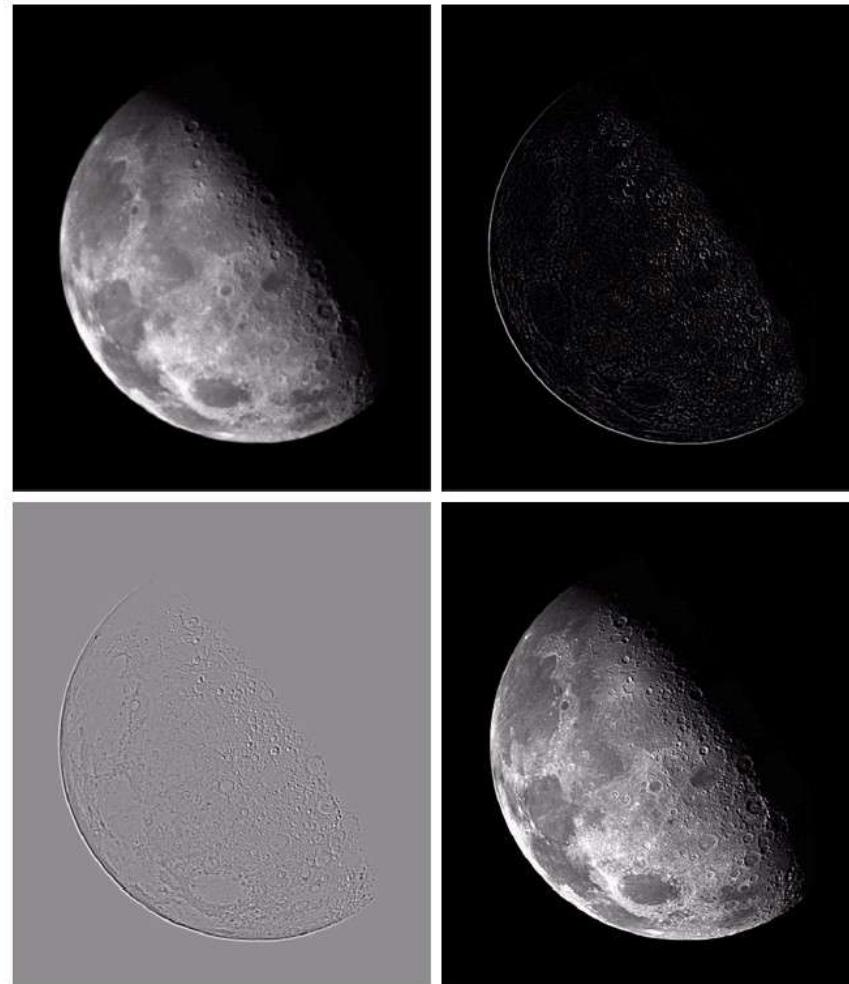
第3章 图像增强

3.2 空域增强

a
b
c
d

FIGURE 3.40

- (a) Image of the North Pole of the moon.
- (b) Laplacian-filtered image.
- (c) Laplacian image scaled for display purposes.
- (d) Image enhanced by using Eq. (3.7-5).
(Original image courtesy of NASA.)



图片来自冈萨雷斯的数字图像处理(英文版)教材

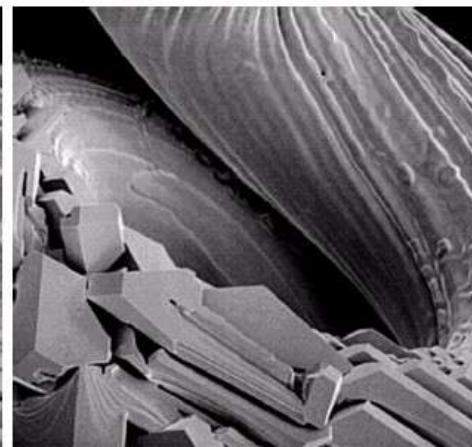
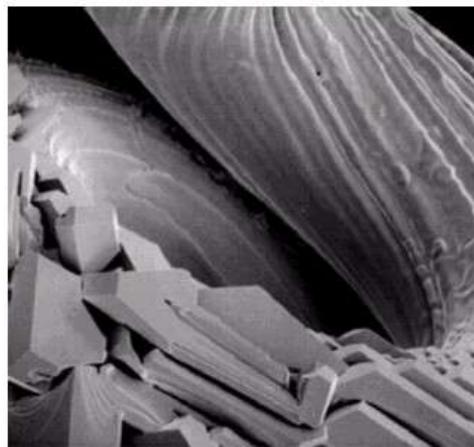
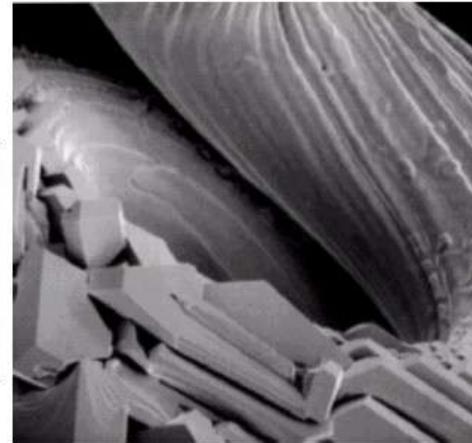


第3章 图像增强

3.2 空域增强

0	-1	0
-1	5	-1
0	-1	0

-1	-1	-1
-1	9	-1
-1	-1	-1



a b c
d e

FIGURE 3.41 (a) Composite Laplacian mask. (b) A second composite mask. (c) Scanning electron microscope image. (d) and (e) Results of filtering with the masks in (a) and (b), respectively. Note how much sharper (e) is than (d). (Original image courtesy of Mr. Michael Shaffer, Department of Geological Sciences, University of Oregon, Eugene.)



数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第3章 图像增强

第6讲 空域增强综合应用

演示文稿说明：

- 本讲内容以板书为主，*ppt* 演示为辅；
- 本讲部分图片来自冈萨雷斯的数字图像处理（英文版）教材。



第3章 图像增强

3.2 空域增强

图像锐化：

$$\begin{aligned} g &= a \times f_{\text{original}} - f_{\text{lowpass}} \\ &= (a-1) \times f_{\text{original}} + f_{\text{highpass}} \end{aligned}$$

若 $a=1$ 时， 拉普拉斯图像？

若 a 越大时， 图像是什么效果
？



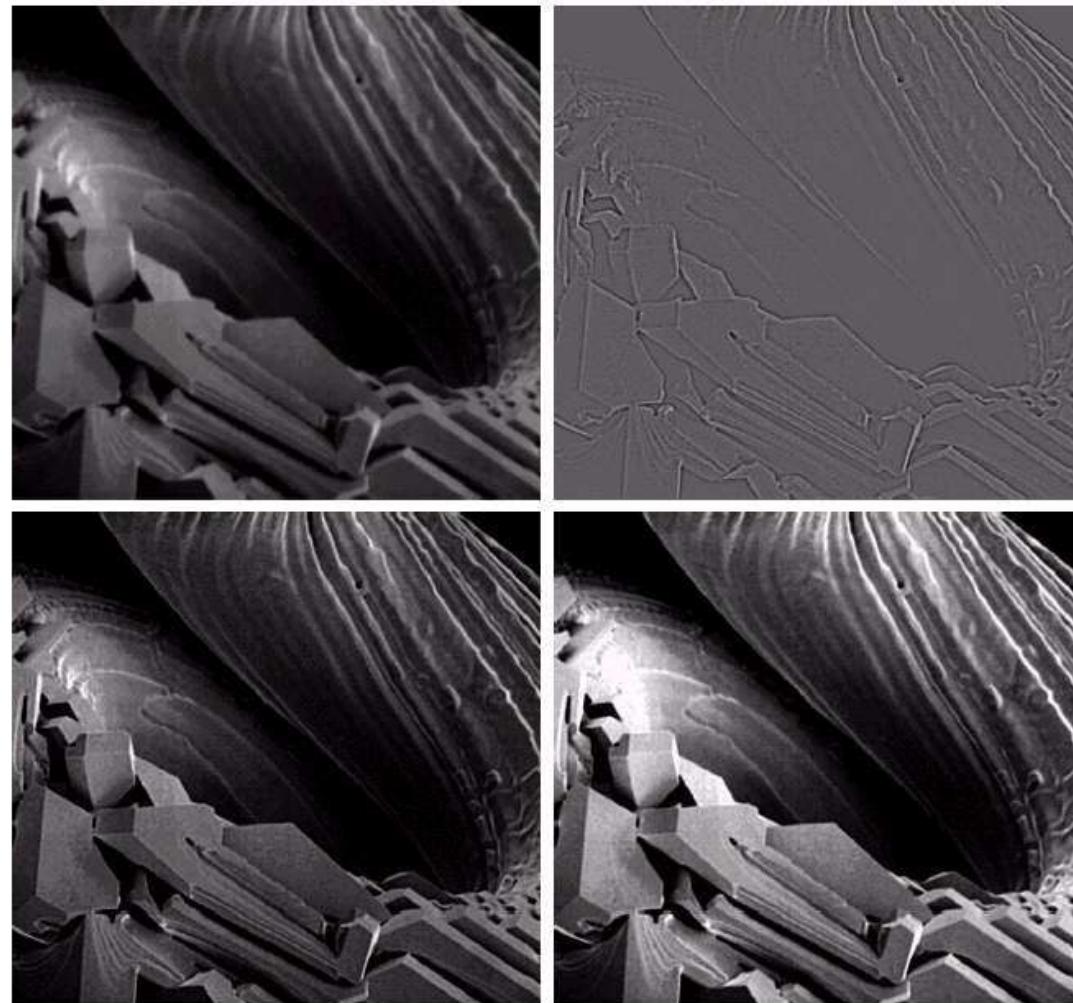
第3章 图像增强

3.2 空域增强

a
b
c
d

FIGURE 3.43

- (a) Same as Fig. 3.41(c), but darker.
(a) Laplacian of (a) computed with the mask in Fig. 3.42(b) using $A = 0$.
(c) Laplacian enhanced image using the mask in Fig. 3.42(b) with $A = 1$. (d) Same as (c), but using $A = 1.7$.

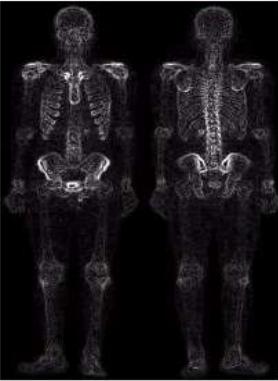
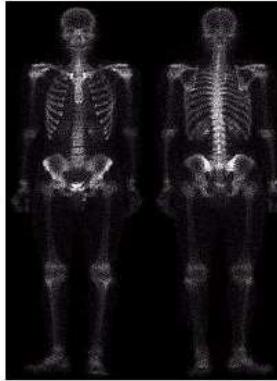
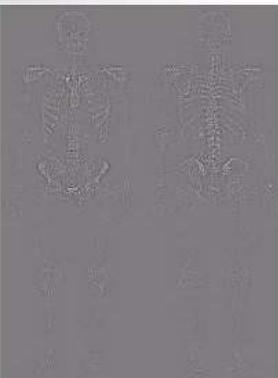
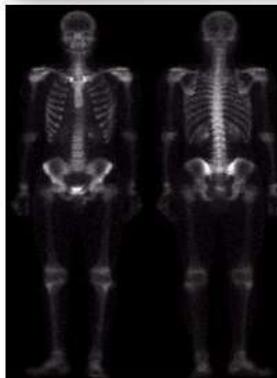


图片来自冈萨雷斯的数字图像处理(英文版)教材



第3章 图像增强

3.2 空域增强

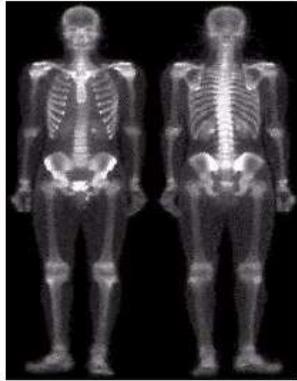
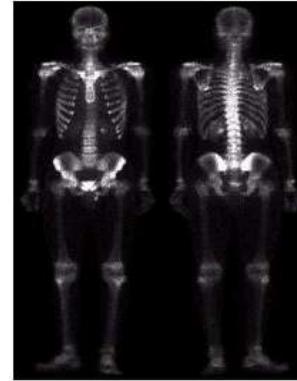
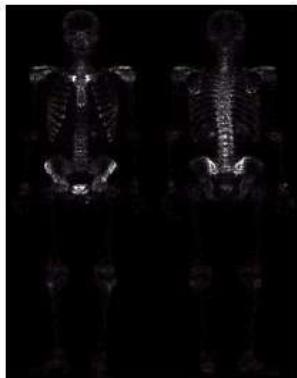
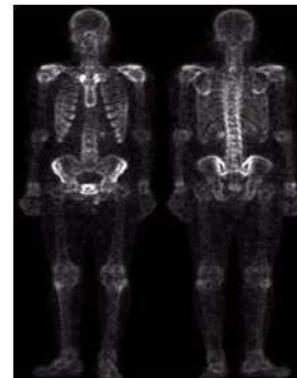


(a) 原图

(b) 拉氏算子锐化

(c) $a+b$

(d) sobel 算子锐化



(e) (d) 的平滑图像

(f) 将 (e) 与 (b)
相乘

(g) (a)+ (f)

(h) 伽玛灰度变化增强
(英文版)

图片来自冈萨雷斯的数字图像处理(英文版)教材



第3章 图像增强

3.2 空域增强

空域增强方法应用举例：

第一步：

拉氏算子

$$f(x, y) + \nabla^2 f$$

-1	-1	-1
-1	8	-1
-1	-1	-1



第3章 图像增强

3.2 空域增强

第二步：

3x3 的 Sobel 算子

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

5x5 的平滑算子

1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1



第3章 图像增强

3.2 空域增强

第三步：动态范围扩展

幂函数



第3章 图像增强

3.3 频域增强

假定原图像 $f(x,y)$ ，经傅立叶变换为 $F(u,v)$ ，频域增强就是选择合适的滤波器函数 $H(u,v)$ 对 $F(u,v)$ 的频谱成分进行调整，然后经逆傅立叶变换得到增强的图像 $g(x,y)$ 。该过程可以通过下面流程描述：

$$f(x,y) \xrightarrow{DFT} F(u,v) \xrightarrow{\text{滤波}} H(u,v) \xrightarrow{IDFT} g(x,y)$$

其中， $G(u, v) = H(u, v) \cdot F(u, v)$, $H(u, v)$ 称为传递函数或滤波器函数。



第3章 图像增强

3.3 频域增强

基本步骤：

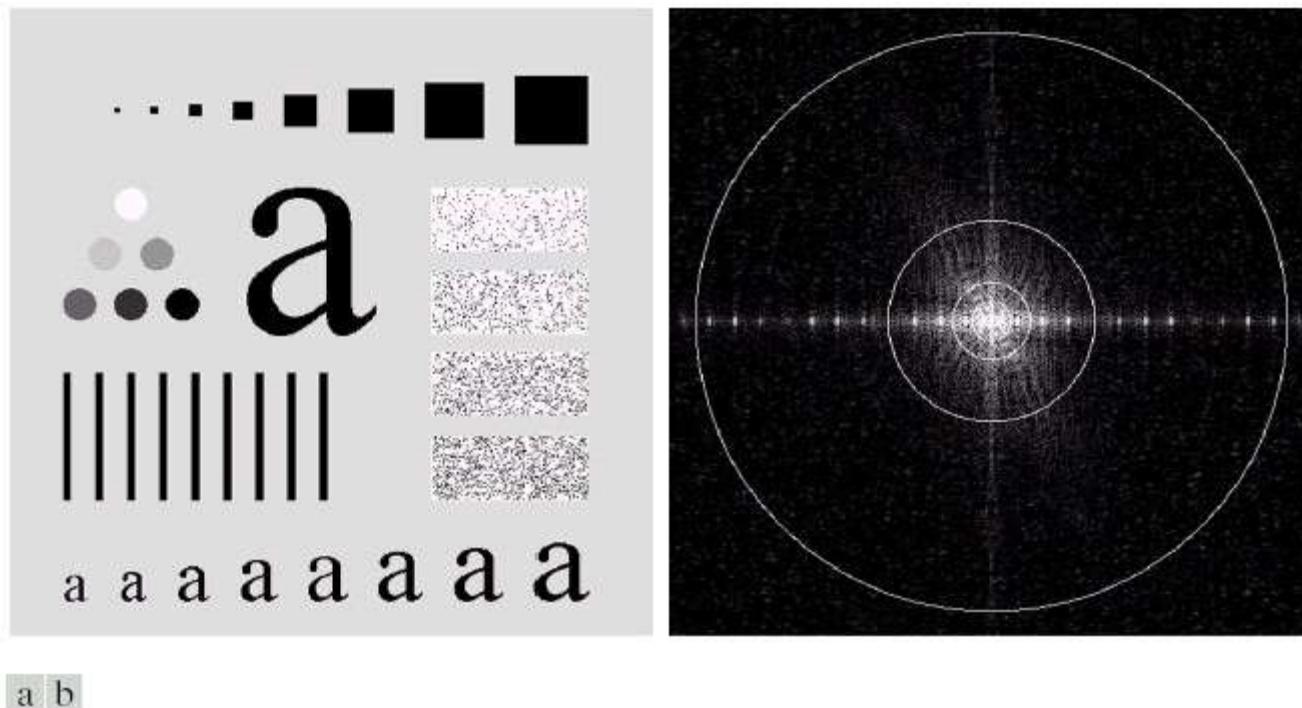
- 将原图像 f 乘以 $(-1)^{x+y}$
- 计算傅立叶变换得 $F(u,v)$
- 选择一个滤波函数 $H(u,v)$ 并乘以 $F(u,v)$
- 进行反傅立叶变换，并取其实部
- 乘以 $(-1)^{x+y}$ 得最后增强的图像

具体得到什么样的滤波结果完全取决于滤波函数 $H(u,v)$



第3章 图像增强

3.3 频域增强



a b

FIGURE 4.11 (a) An image of size 500×500 pixels and (b) its Fourier spectrum. The superimposed circles have radii values of 5, 15, 30, 80, and 230, which enclose 92.0, 94.6, 96.4, 98.0, and 99.5% of the image power, respectively.

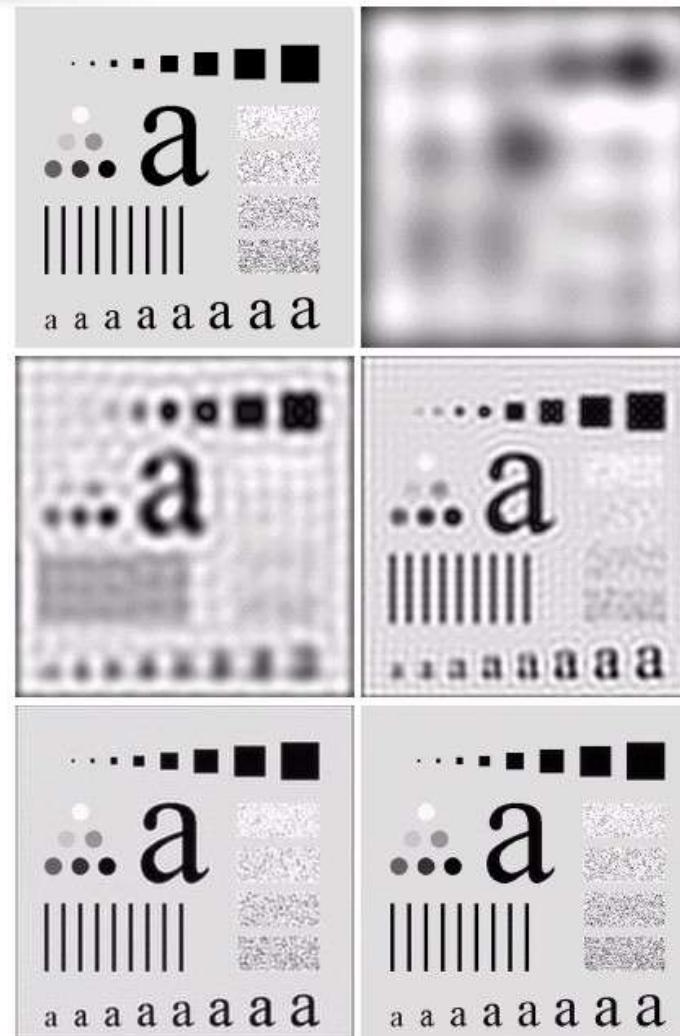
图片来自冈萨雷斯的数字图像处理(英文版)教材



第3章 图像增强

3.3 频域增强

- ✓ 低通滤波



图片来自冈萨雷斯的数字图像处理(英文版)教材



数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第3章 图像增强

第7讲 频域增强

演示文稿说明：

- 本讲内容以板书为主，*ppt* 演示为辅；
- 本讲部分图片来自冈萨雷斯的数字图像处理（英文版）教材。



第3章 图像增强

3.3 频域增强

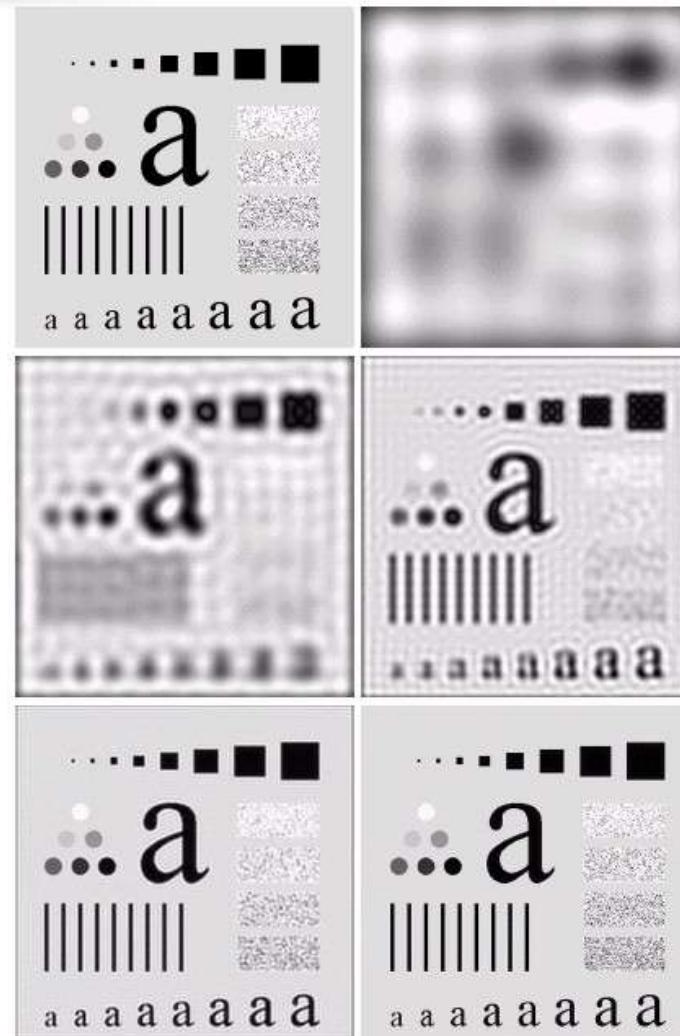
- ✓ 低通滤波
 - 理想低通滤波 (ILPF)
 - Butterworth 低通滤波 (BLPF)
 - Gaussian 低通滤波 (GLPF)



第3章 图像增强

3.3 频域增强

- ✓ 低通滤波



图片来自冈萨雷斯的数字图像处理(英文版)教材



第3章 图像增强

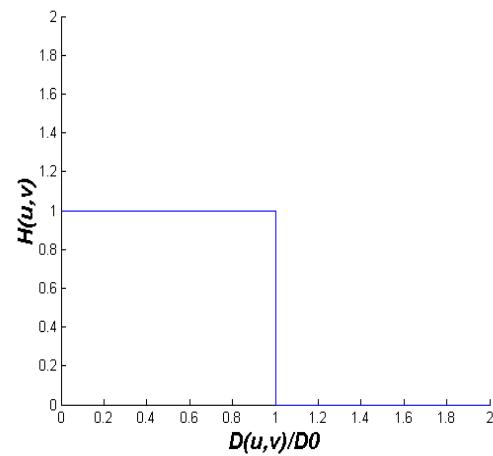
3.3 频域增强

➤ 理想低通滤波 (ILPF)

$G(u, v) = H(u, v) \cdot F(u, v)$ $H(u, v)$ 称为传递函数或滤波器函数。

$$g(x, y) = h(u, v) * f(x, y)$$

$$H(u, v) = \begin{cases} 1 & D(u, v) \leq D_0 \\ 0 & D(u, v) > D_0 \end{cases}$$





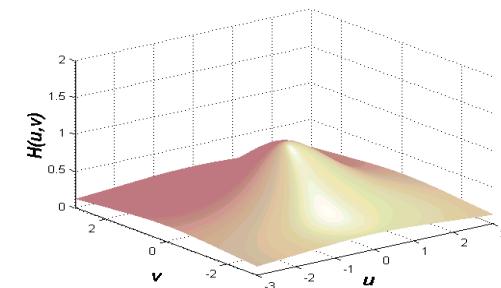
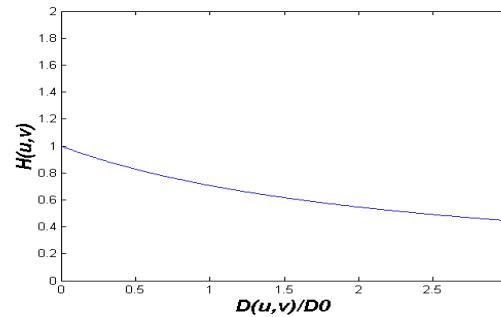
第3章 图像增强

3.3 频域增强

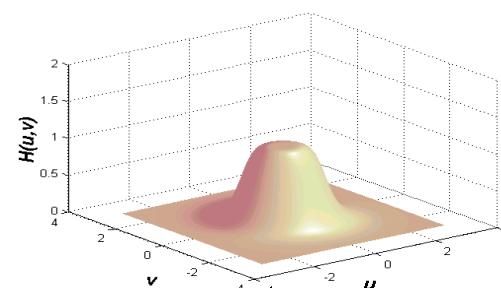
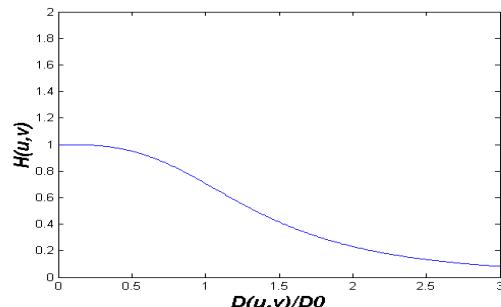
- Butterworth 低通滤波 (BLPF)

$$H(u, v) = \frac{1}{1 + k[D(u, v)/D_0]^{2n}}$$

n=1



n=3





第3章 图像增强

3.3 频域增强

➤ BLPF

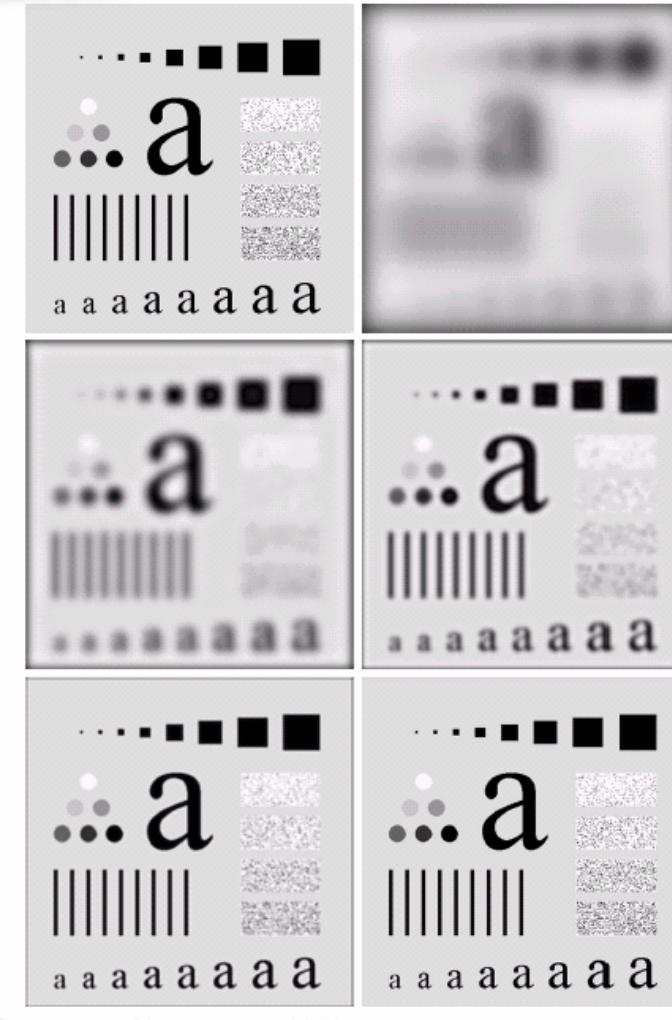


FIGURE 4.15 (a) Original image. (b)–(f) Results of filtering with BLPFs of order 2, with cutoff frequencies at radii of 5, 15, 30, 80, and 230, as shown in Fig. 4.11(b). Compare with Fig. 4.12.

图片来自冈萨雷斯的数字图像处理（英文版）教材



第3章 图像增强

3.3 频域增强

- Gaussian 低通滤波 (GLPF)

$$H(u, v) = e^{-D^2(u, v)/2\sigma^2}$$



第3章 图像增强

3.3 频域增强

➤ GLPF

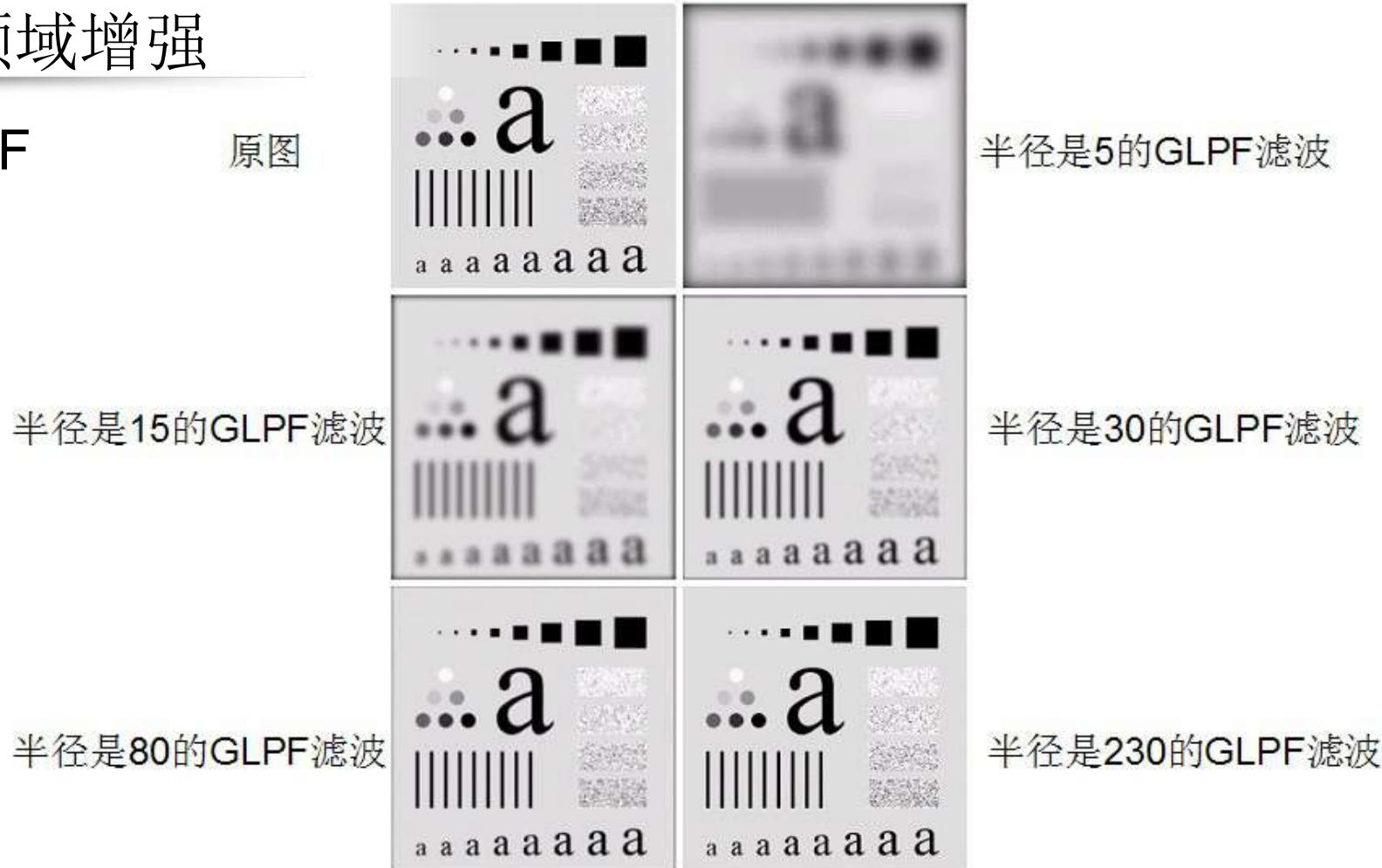


FIGURE 4.18 (a) Original image. (b)-(f) Results of filtering with Gaussian lowpass filters with cutoff frequencies set at radii values of 5, 15, 30, 80, and 230, as shown in Fig. 4.11(b). Compare with Figs. 4.12 and 4.15.

a
b
c
d
e
f

图片来自冈萨雷斯的数字图像处理(英文版)教材



第3章 图像增强

3.3 频域增强

- ✓ 高通滤波

$$H_{hp}(u, v) = 1 - H_{lp}(u, v)$$

- 理想高通滤波 (IHPF)
- Butterworth 高通滤波 (BHPF)
- Gaussian 高通滤波 (GHPF)

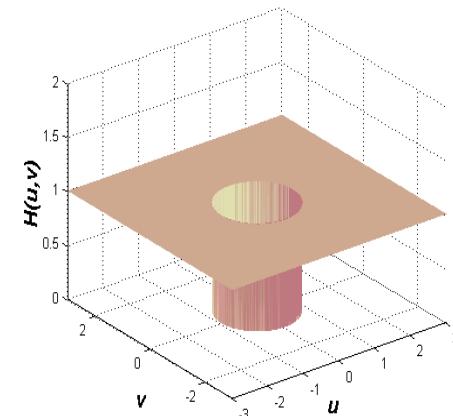
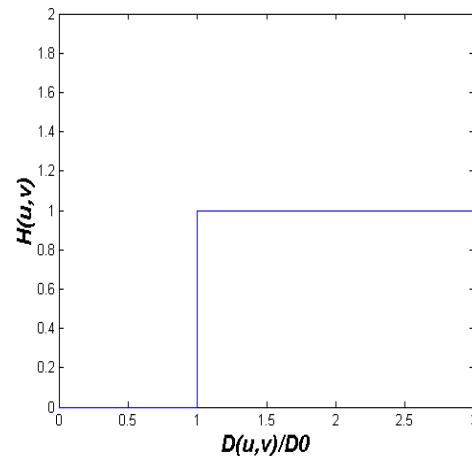


第3章 图像增强

3.3 频域增强

- 理想高通滤波 (IHPF)

$$H(u, v) = \begin{cases} 0 & D(u, v) \leq D_0 \\ 1 & D(u, v) > D_0 \end{cases}$$





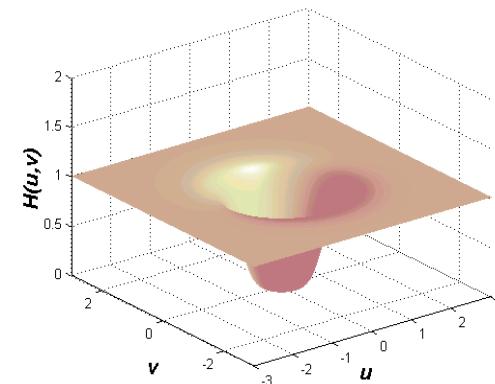
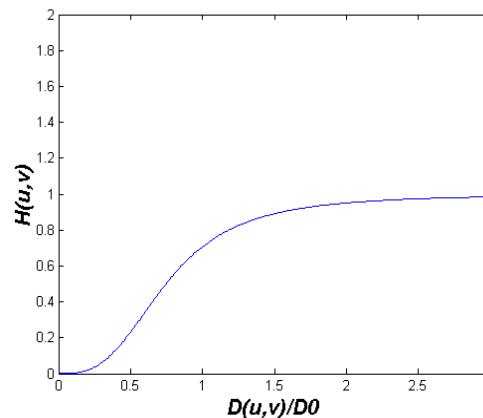
第3章 图像增强

3.3 频域增强

- Butterworth 高通滤波 (BHPF)

$$H(u, v) = \frac{1}{1 + [D_0/D(u, v)]^{2n}}$$

n=3

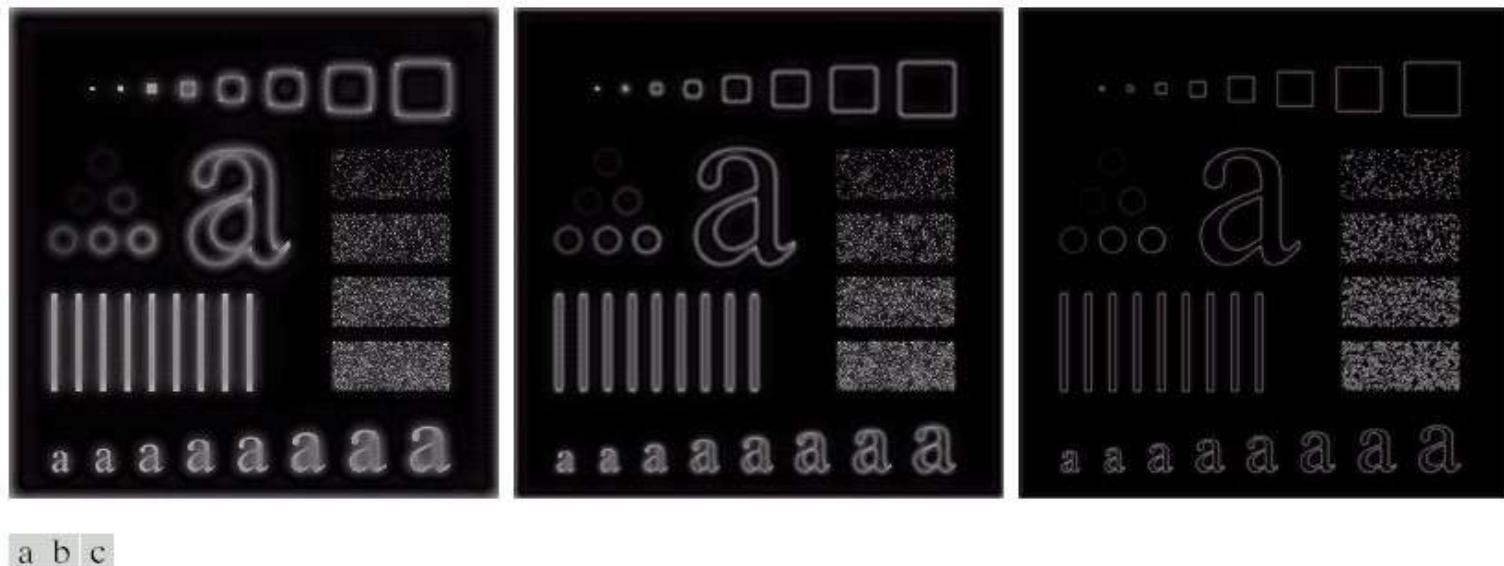




第3章 图像增强

3.3 频域增强

➤ BHPF



a b c

FIGURE 4.25 Results of highpass filtering the image in Fig. 4.11(a) using a BHPF of order 2 with $D_0 = 15$, 30, and 80, respectively. These results are much smoother than those obtained with an ILPF.

图片来自冈萨雷斯的数字图像处理(英文版)教材

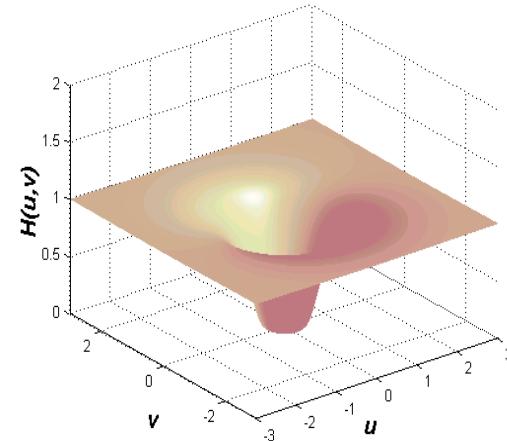
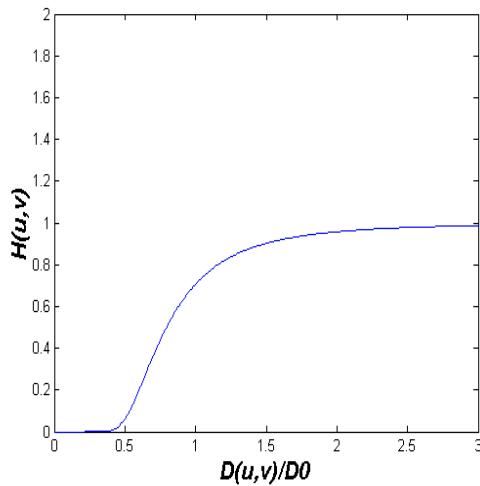


第3章 图像增强

3.3 频域增强

- Gaussian 高通滤波 (GHPF)

$$H(u, v) = 1 - e^{-[D^2(u, v)/2D^2_0]}\}$$





第3章 图像增强

3.3 频域增强

➤ GHPF

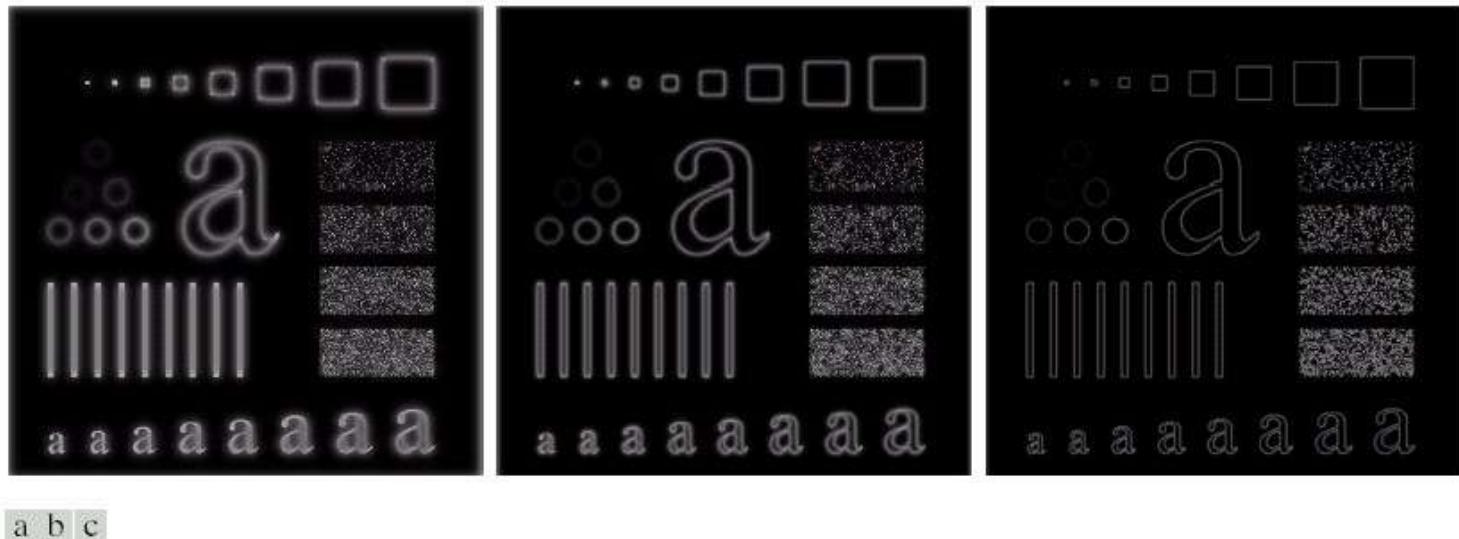


FIGURE 4.26 Results of highpass filtering the image of Fig. 4.11(a) using a GHPF of order 2 with $D_0 = 15$, 30, and 80, respectively. Compare with Figs. 4.24 and 4.25.

图片来自冈萨雷斯的数字图像处理(英文版)教材



第3章 图像增强

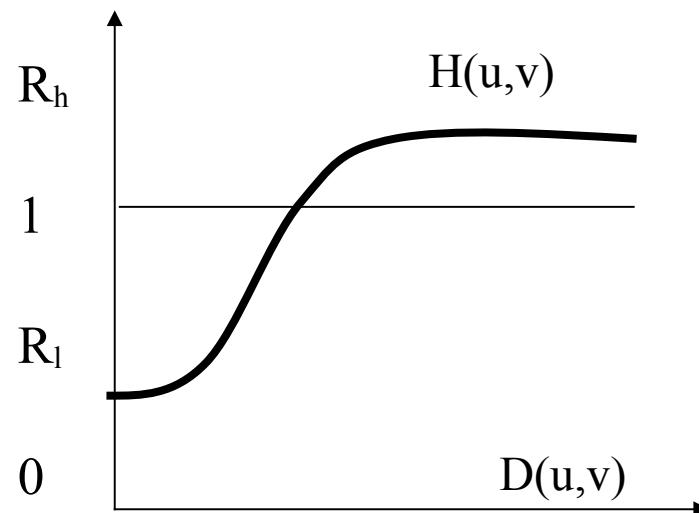
3.3 频域增强

✓ 同态滤波

同态滤波是基于成像模型：

$$f(x, y) = i(x, y)r(x, y)$$

在这个模型中，入射光部分对应于图像中的低频部分，反射光部分对应于图像中的高频部分。下图即是同态滤波过程及其滤波器。





第3章 图像增强

3.3 频域增强

- ✓ 同态滤波



a. 原图



b. 同态滤波后的图



第3章 图像增强

3.3 频域增强

- ✓ 同态滤波



a. 原图



b. 同态滤波后的图像



数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第3章 图像增强

第8讲 彩色增强

演示文稿说明：

本讲部分图片来自冈萨雷斯的数字图像处理（英文版）教材。



3.4 彩色增强

彩色增强

1. 彩色基础知识
2. 彩色模型
3. 伪彩色增强
4. 真彩色增强



3.4 彩色增强

1. 彩色基础知识

➤ 三原色

红色（Red）、绿色（Green）、蓝色（Blue）

➤ 原色相加可产生二次色

深红色：红+蓝

青色： 绿+蓝

黄色： 红+绿



3.4 彩色增强

2. 彩色模型（或颜色空间）

➤ RGB

- CCD 技术直接感知 R,G,B 三个分量
- 是图像成像、显示、打印等设备的基础

➤ HSI (色调、饱和度、亮度)

- 两个特点：
 - I 分量与图像的彩色信息无关
 - H 和 S 分量与人感受颜色的方式是紧密相连的
- 将亮度 (I) 与色调 (H) 和饱和度 (S) 分开
- 避免颜色受到光照明暗 (I) 等条件的干扰
- 仅仅分析反映色彩本质的色调和饱和度
- 广泛用于计算机视觉、图像检索和视频检索



3.4 彩色增强

➤ YUV

- Y 指亮度 (Brightness), 即灰度值
- U 和 V 指色调
- 用于彩色电视广播, 被欧洲的电视系统所采用 (属于 PAL 系统)
- Y 分量可提供黑白电视机的所有影像信息

➤ YCbCr

- Y 指亮度, 与 YUV 的 Y 相同
- Cb 和 Cr 由 U 和 V 调整得到
- JPEG 采用的彩色空间



3.4 彩色增强

彩色空间转换：

- a. RGB → HSI
- b. RGB → YUV
- c. RGB → YCbCr



3. 4 彩色增强

a. RGB → HSI

$$H = \begin{cases} \theta & B \leq G \\ 360 - \theta & B > G \end{cases}$$

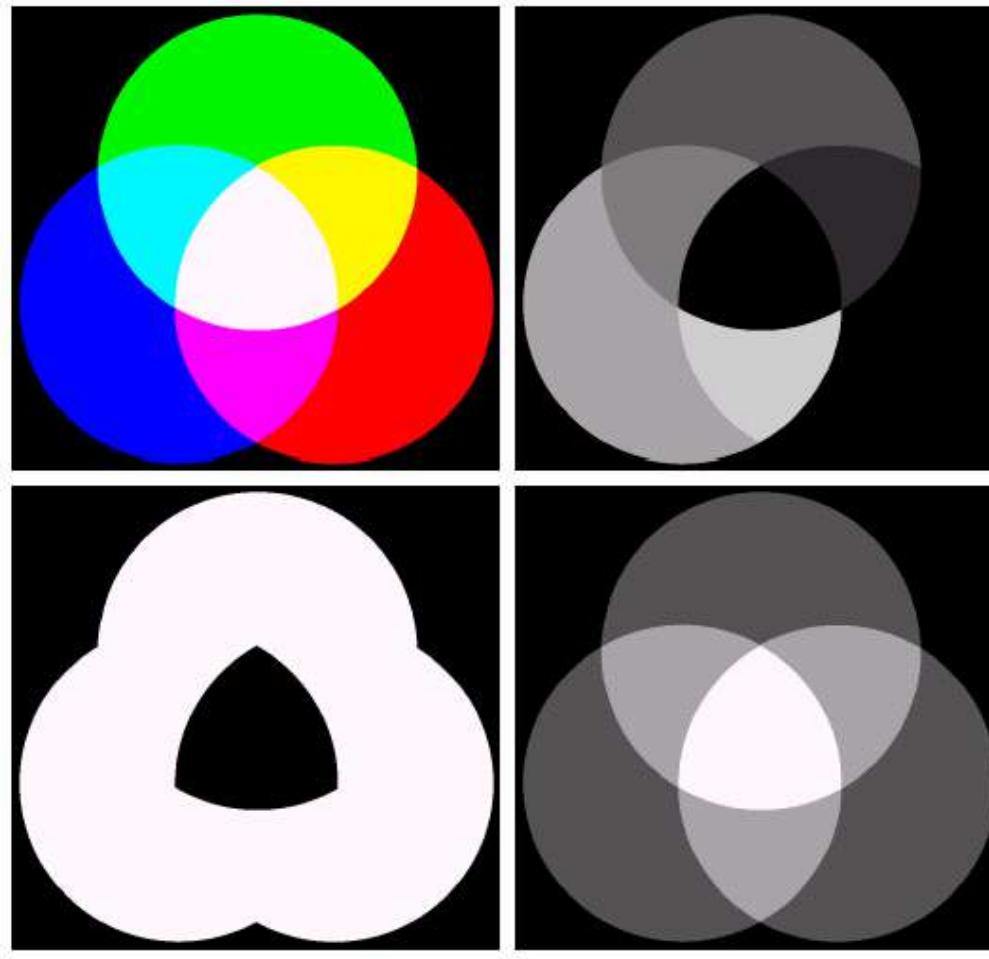
$$\theta = \arccos \left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2 + (R-G)(G-B)]^{\frac{1}{2}}} \right\}$$

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)]$$

$$I = \frac{1}{3}(R + G + B)$$



3.4 彩色增强



a b
c d

FIGURE 6.16 (a) RGB image and the components of its corresponding HSI image:
(b) hue, (c) saturation, and (d) intensity.



3. 4 彩色增强

HSI → RGB

$$0^0 \leq H < 120^0$$

$$R = I \left[1 + \frac{S \cos H}{\cos(60^0 - H)} \right] \quad G = 1 - (R+B) \quad B = I(1-s)$$

$$120^0 \leq H < 240^0 \quad H = H - 120^0$$

$$G = I \left[1 + \frac{S \cos H}{\cos(60^0 - H)} \right] \quad B = 1 - (R+G) \quad R = I(1-s)$$

$$240^0 \leq H < 360^0 \quad H = H - 240^0$$

$$B = I \left[1 + \frac{S \cos H}{\cos(60^0 - H)} \right] \quad R = 1 - (G+B) \quad G = I(1-s)$$



3.4 彩色增强

b. RGB → YUV

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.148 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.140 \\ 1 & -0.395 & -0.581 \\ 1 & 2.032 & 0 \end{bmatrix} \begin{bmatrix} Y \\ U \\ V \end{bmatrix}$$



3.4 彩色增强

c. RGB → YCbCr

$$\begin{bmatrix} Y \\ Cb \\ Cr \\ 1 \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 & 0 \\ -0.1687 & -0.3313 & 0.5000 & 128 \\ 0.5000 & -0.4187 & -0.0813 & 128 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 1.40200 & 0 \\ 1 & -0.34414 & -0.71414 \\ 1 & 1.77200 & 0 \end{bmatrix} \begin{bmatrix} Y \\ Cb - 128 \\ Cr - 128 \end{bmatrix}$$



3.4 彩色增强

3. 伪彩色图像处理

定义：或假彩色图像处理， 根据一定的准则对灰度值赋以彩色的处理。

分类：伪彩色图像、真彩色图像、单色图像。

作用：人类可以辨别上千种颜色和强度，但只能辨别二十几种灰度。

伪彩色图像处理相关技术

- a. 强度分层技术
- b. 灰度级映射到彩色技术



3.4 彩色增强

a. 强度分层技术

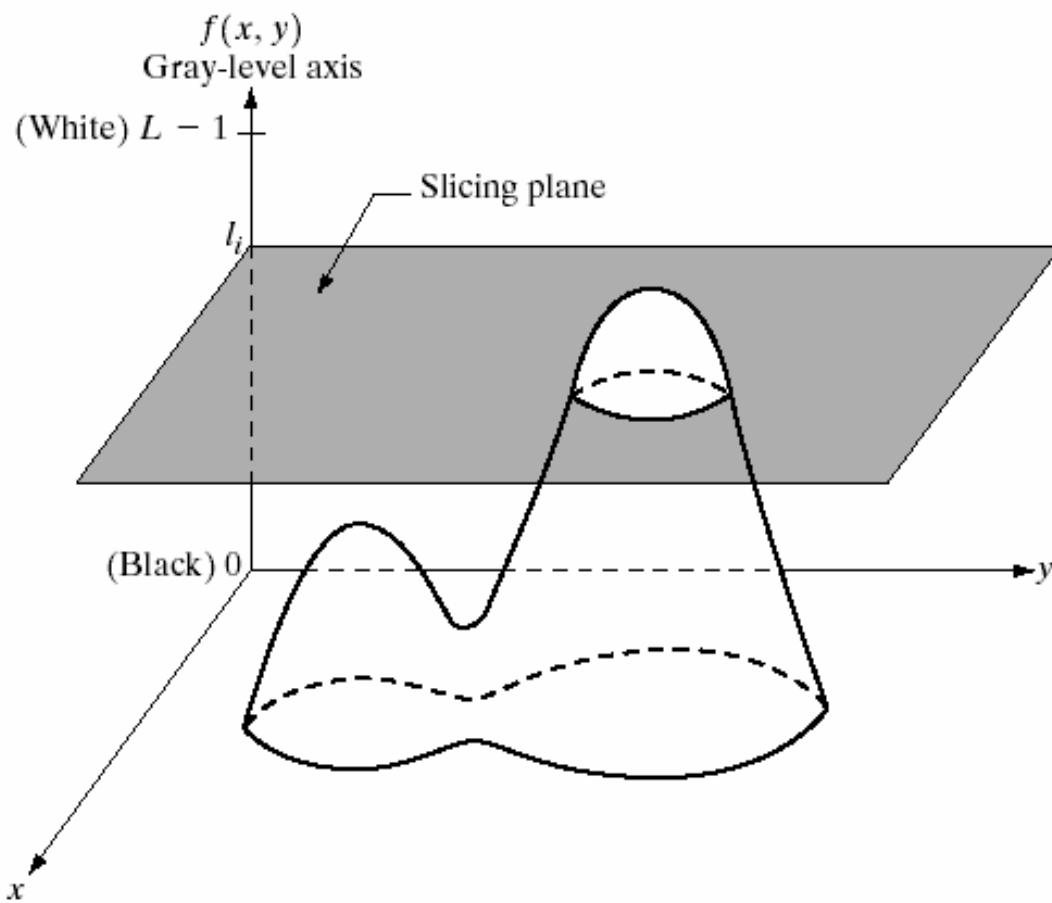


FIGURE 6.18 Geometric interpretation of the intensity-slicing technique.



3.4 彩色增强

强度分层技术定义：

令 $[0, L-1]$ 表示灰度级，使 l_0 代表黑色 ($f(x, y) = 0$)， l_{L-1} 代表白色 ($f(x, y) = L-1$)。假设垂直于强度轴的P个平面定义为量级 l_1, l_2, \dots, l_P 。 $0 < P < L-1$ ，P个平面将灰度级分为 $P+1$ 个间隔， V_1, V_2, \dots, V_{P+1} ，则灰度级到彩色的赋值关系：

$$f(x, y) = c_k \quad f(x, y) \in V_k$$

c_k 是与强度间隔 V_k 第K级强度有关的颜色

V_k 是由在 $l=k-1$ 和 $l=k$ 分割平面定义的



3.4 彩色增强

强度分层技术：

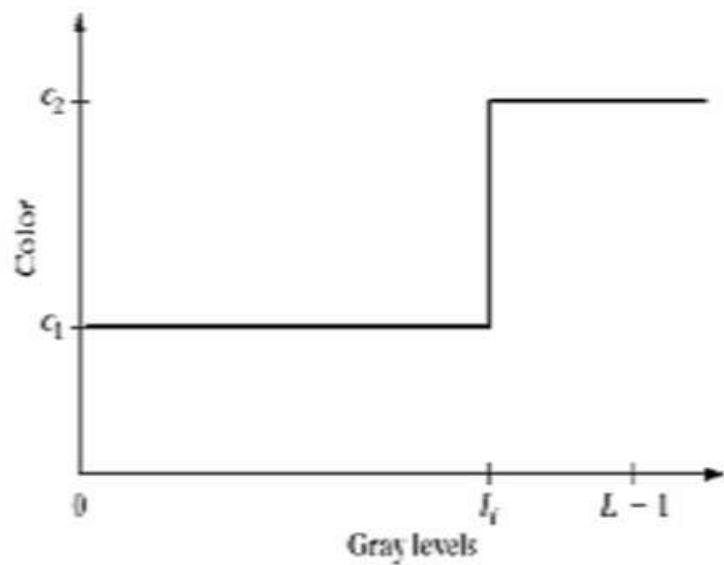
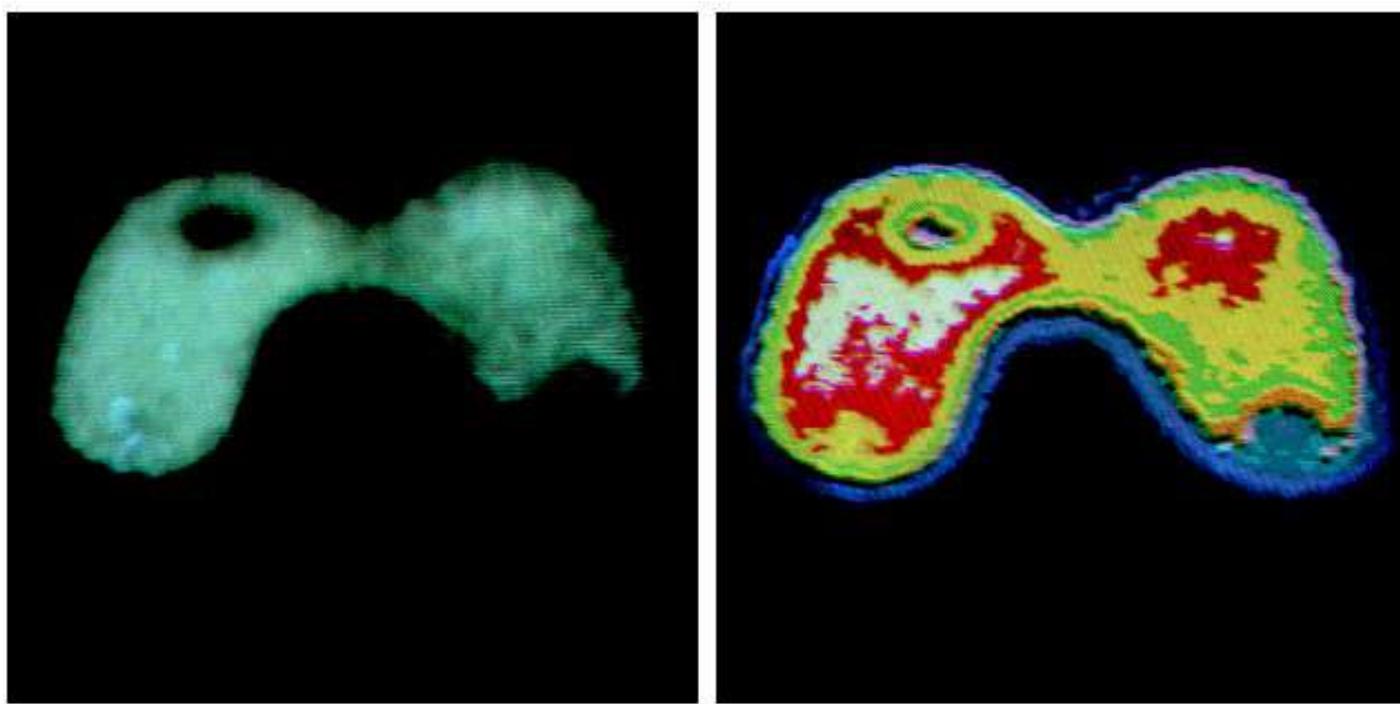


FIGURE 6.19 An alternative representation of the intensity-slicing technique.



3.4 彩色增强

举例 1：甲状腺模型



a b

FIGURE 6.20 (a) Monochrome image of the Picker Thyroid Phantom. (b) Result of density slicing into eight colors. (Courtesy of Dr. J. L. Blankenship, Instrumentation and Controls Division, Oak Ridge National Laboratory.)

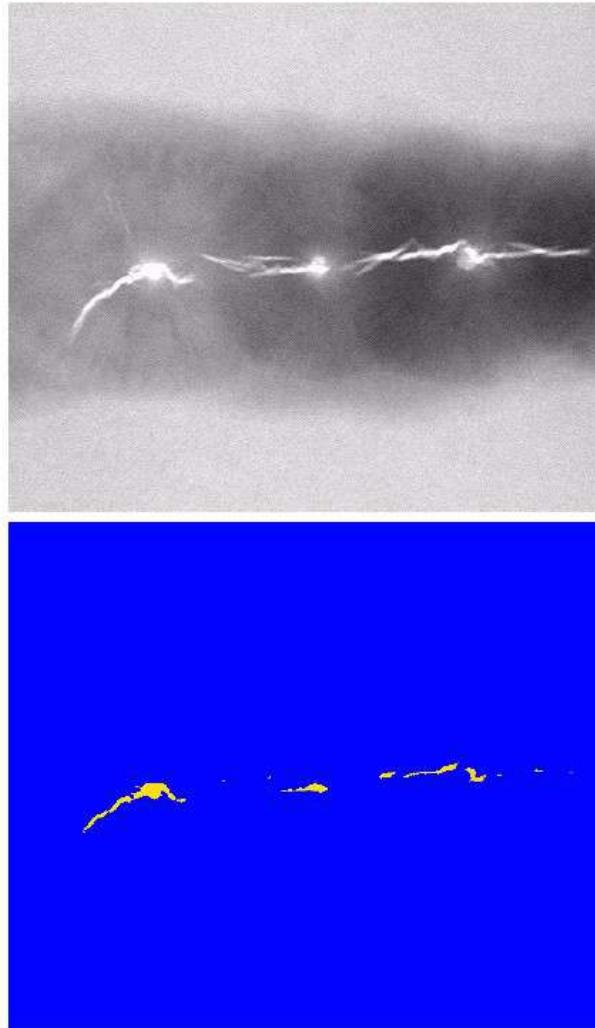


3.4 彩色增强

举例 2：焊点问题检测

a
b

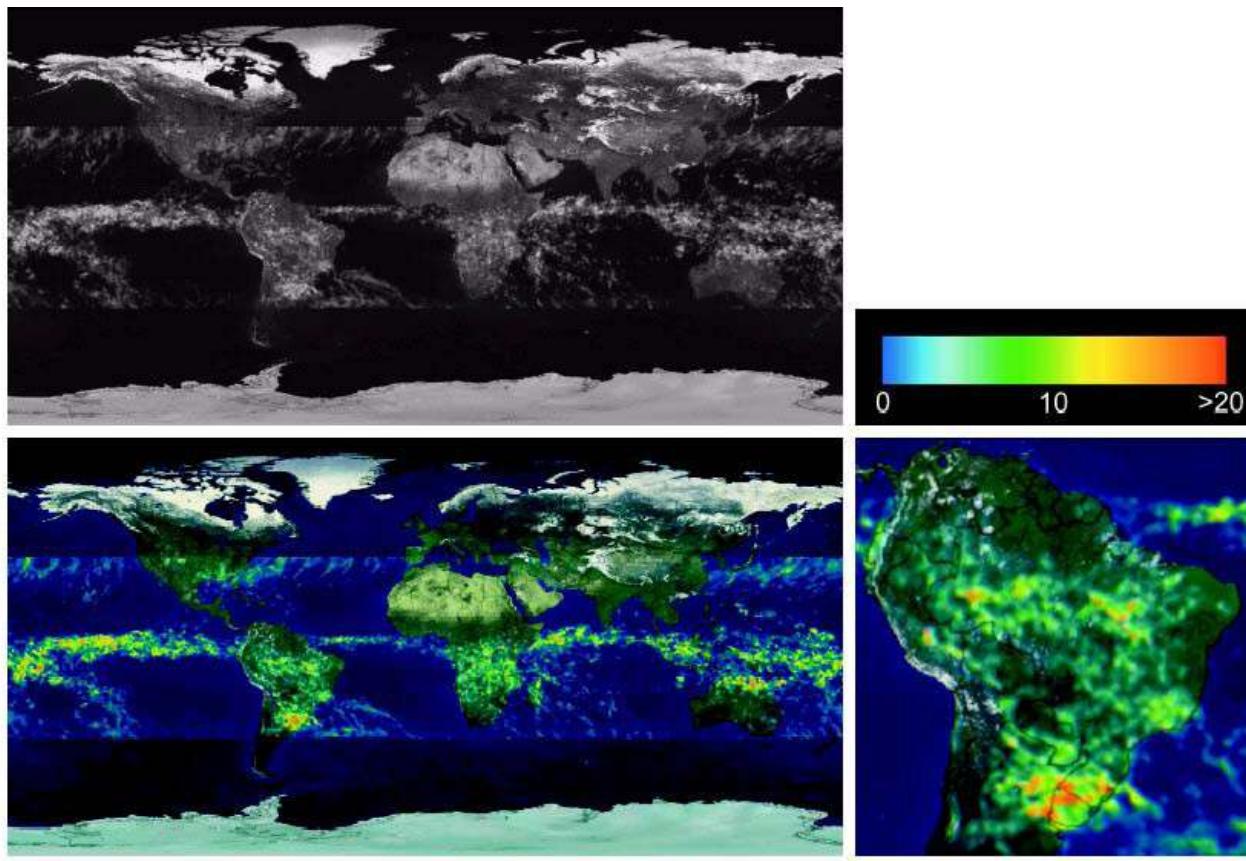
FIGURE 6.21
(a) Monochrome X-ray image of a weld. (b) Result of color coding.
(Original image courtesy of X-TEK Systems, Ltd.)





3.4 彩色增强

举例 3：降雨量问题



a b
c d

FIGURE 6.22 (a) Gray-scale image in which intensity (in the lighter horizontal band shown) corresponds to average monthly rainfall. (b) Colors assigned to intensity values. (c) Color-coded image. (d) Zoom of the South America region. (Courtesy of NASA.)



3.4 彩色增强

b. 灰度级映射到彩色技术

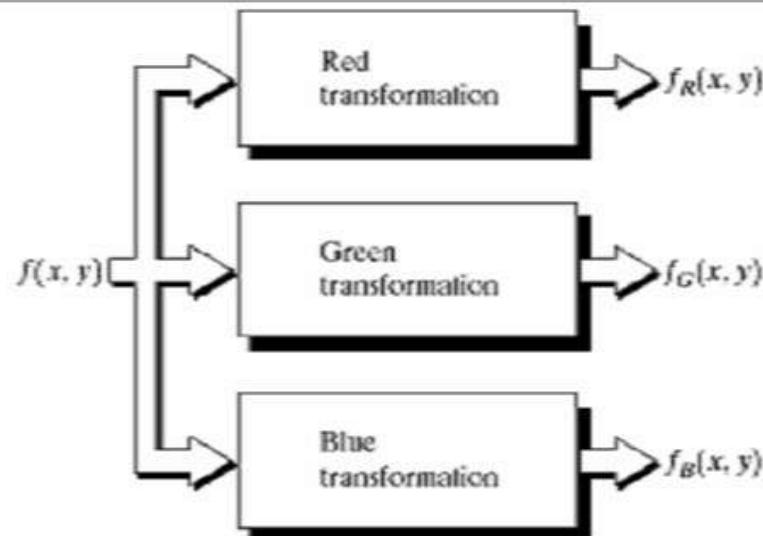
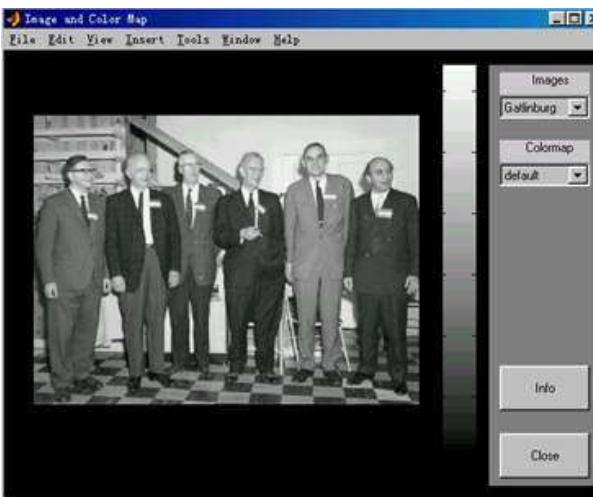
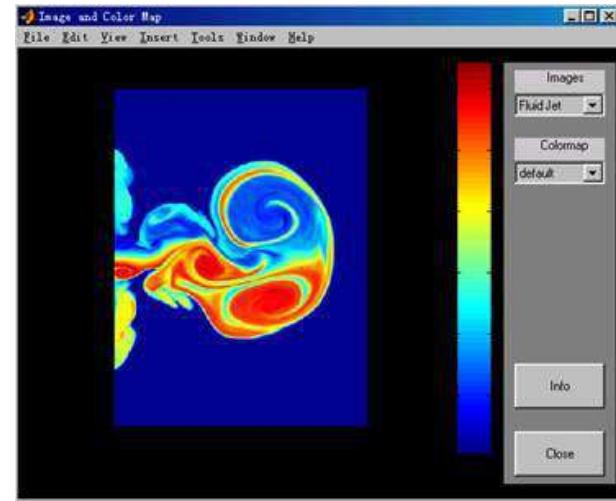


FIGURE 6.23 Functional block diagram for pseudocolor image processing. f_R , f_G , and f_B are fed into the corresponding red, green, and blue inputs of an RGB color monitor.

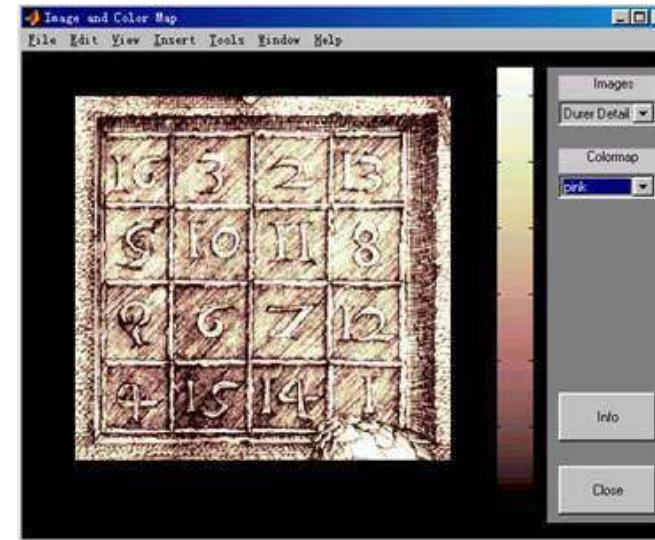


3.4 彩色增强



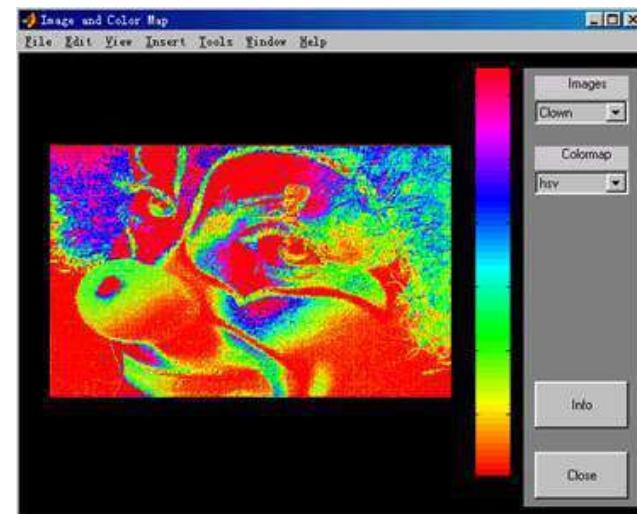


3.4 彩色增强





3.4 彩色增强





3. 4 彩色增强

4. 真彩色增强

- a. 彩色图像平滑
- b. 彩色图像锐化



3.4 彩色增强

a. 彩色图像平滑

令 S_{xy} 表示在RGB彩色图像中定义一个中心在 (x, y) 的邻域的坐标集，在该邻域中RGB分量的平均值为

$$\bar{c}(x, y) = \frac{1}{k} \sum_{(x, y) \in S_{xy}} c(x, y)$$

$$\bar{c}(x, y) = \begin{bmatrix} \frac{1}{K} \sum_{(x, y) \in S_{xy}} R(x, y) \\ \frac{1}{K} \sum_{(x, y) \in S_{xy}} G(x, y) \\ \frac{1}{K} \sum_{(x, y) \in S_{xy}} B(x, y) \end{bmatrix}$$



3.4 彩色增强

彩色图像平滑举例 1：



a
b
c
d

FIGURE 6.38
(a) RGB image.
(b) Red
component image.
(c) Green
component.
(d) Blue
component.



3.4 彩色增强

彩色图像平滑举例 2：



a b c

FIGURE 6.39 HSI components of the RGB color image in Fig. 6.38(a). (a) Hue. (b) Saturation. (c) Intensity.



3.4 彩色增强

彩色图像平滑举例 3：



a b c

FIGURE 6.40 Image smoothing with a 5×5 averaging mask. (a) Result of processing each RGB component image. (b) Result of processing the intensity component of the HSI image and converting to RGB. (c) Difference between the two results.



3.4 彩色增强

b. 彩色图像锐化

RGB彩色空间，分别计算每一分量图像的拉普拉斯变换

$$\nabla^2 [c(x, y)] = \begin{bmatrix} \nabla^2 R(x, y) \\ \nabla^2 G(x, y) \\ \nabla^2 B(x, y) \end{bmatrix}$$

$$\begin{aligned} g(x, y) &= f(x, y) - \nabla^2 f(x, y) \\ &= f(x, y) - \{[f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1)] - 4f(x, y)\} \\ &= 5f(x, y) - [f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1)] \end{aligned}$$



3.4 彩色增强

彩色图像锐化举例 1：



a b c

FIGURE 6.41 Image sharpening with the Laplacian. (a) Result of processing each RGB channel. (b) Result of processing the intensity component and converting to RGB. (c) Difference between the two results.

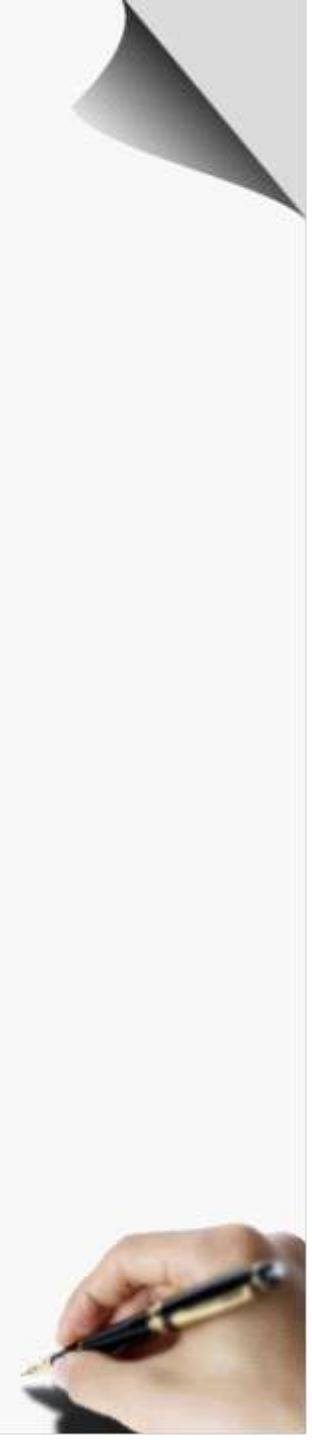


数字图像处理

(Digital Image Processing)

山东科技大学信息与电气工程学院

孙农亮





本章内容

- 第 1 讲 图像复原的概念、退化原因、线性移不变系统
- 第 2 讲 图像退化的数学模型及表达、常见退化模型
- 第 3 讲 无约束复原 - 逆滤波方法
- 第 4 讲 有约束复原 - 维纳滤波和约束最小平方滤波

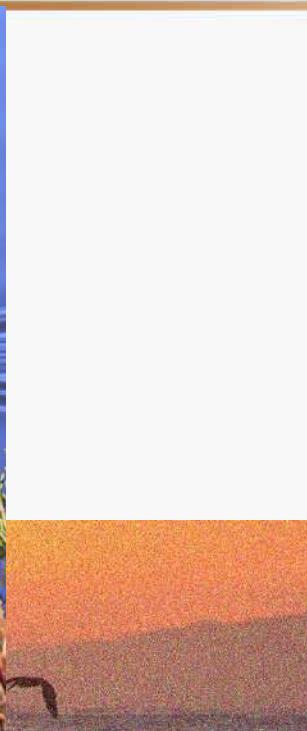
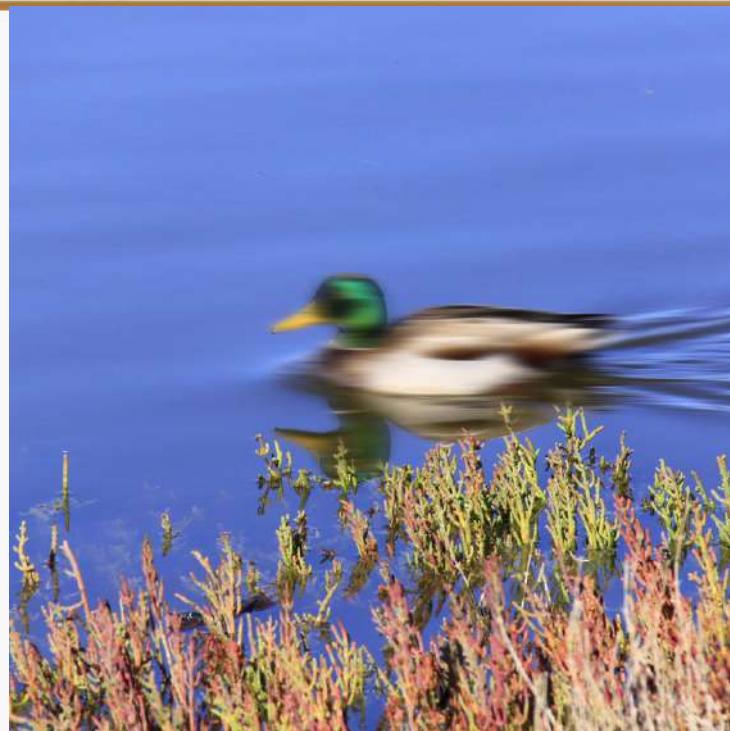


第4章 图像复原

第1讲 图像复原的概念、退化原因、线性移不变系统



第4章 图像复原





4.1 引言

(1) 图像复原的概念

图像复原也称图像恢复，是图像处理中的一大类技术。所谓图像复原，是指通过某些方法、手段、规则从退化了的图像，恢复出原始的图像。

引起退化的因素很多，如图像在形成、传输、记录过程中，由于光学系统调焦不准、相机和物体之间的相对运动，遥感图像中的大气扰动、摄像胶片的非线性和几何畸变、噪声干扰等。

在进行图像复原时，既可以用连续数学，也可以用离散数学进行处理。其次，处理既可在空间域，也可在频域进行。



4.1 引言

(2) 图像恢复与图像增强的异同

相同点：改进输入图像的视觉质量。

不同点：图像增强目的是取得较好的视觉结果（不考虑退化原因）； 图像恢复根据相应的退化模型和知识重建或恢复原始的图像（考虑退化原因）。主观的，客观的？



4.1 引言

(3) 图像退化的原因

图像退化指由场景得到的图像没能完全地反映场景的真实内容，产生了失真等问题。其原因是多方面的。如：

透镜象差 / 色差

聚焦不准（失焦，限制了图像锐度）

模糊（限制频谱宽度）

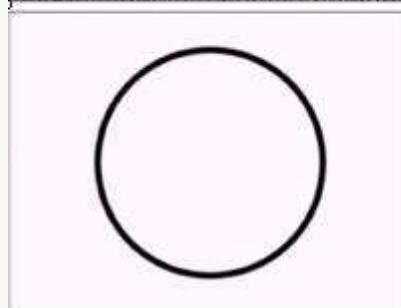
噪声（是一个统计过程）

抖动（机械、电子）等



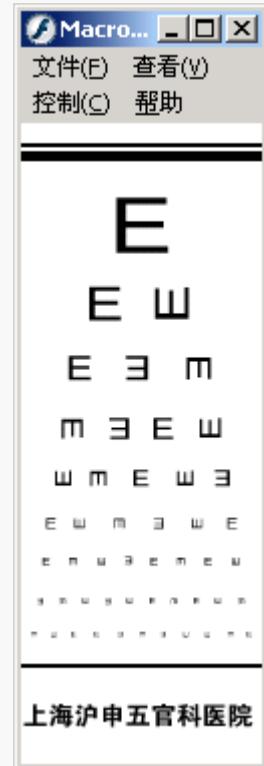
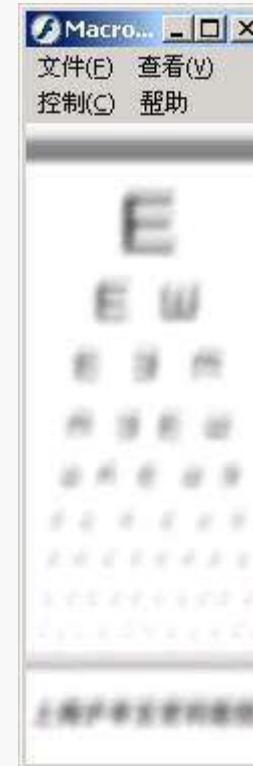
4.1 引言

图像退化举例：



Blur caused by motion

After restoration



上海沪申五官科医院

注：举例图片来自于 Rafael C. Gonzalez 《Digital Image Processing Second Edition》一书和网络



4.1 引言

(4) 噪声及其特性

噪声是最常见的退化因素之一，对信号来说，噪声是一种外部干扰。但噪声本身也是一种信号（携带了噪声源的信息）。

1) 关于噪声的度量

人们常只关心噪声的强度，可用信噪比、能量比（电压平方比）等来描述。分别表示为：

$$SNR = 10 \log_{10} \left(\frac{V_s^2}{V_n^2} \right) \quad SNR = \left(\frac{C_{ob}}{\sigma} \right)^2 = \left[\frac{\text{灰度对比度}}{\text{噪声均方差}} \right]^2$$



4.1 引言

2) 常见噪声

白噪声（频率覆盖整个频谱均匀），高斯噪声（幅度符合高斯分布），粉色噪声（在对数频率间隔内有相同的能量），发射噪声，均匀噪声，椒盐噪声等。



4.2 图像退化的数学模型

(1) 系统 H 的基本定义

就一般而言，系统是某些元件或部件以某种方式构造而成的整体。系统本身所具有的某些特性就构成了通过系统的输入信号与输出信号的某种联系。在数学上，看作某种变换，或叫做算子。用 H 表示

(2) 退化模型示意图

其中 H 为退化过程， $n(x, y)$ 为加性噪声（统计特性已知）。

$$g(x,y)=H[f(x,y)]+n(x,y)$$



4.2 图像退化的数学模型

(3) 线性移不变系统

系统的分类可有：线性系统和非线性系统，时变系统和非时变系统，集总参数系统和分布参数系统，连续系统和离散系统。

1) 线性系统：是具有均匀性和相加性的系统

$$H[a f_1(x,y) + b f_2(x,y)] = a H[f_1(x,y)] + b H[f_2(x,y)]$$

2) 时不变系统：满足各个参数不随时间变化。

3) 线性空间不变系统：满足

$$H[f(x - \alpha, y - \beta)] = g(x - \alpha, y - \beta)$$

对于线性空不变系统，

$$g(x,y) = f(x,y) * h(x,y) + n(x,y)$$

点扩散函数（PSF）





数字图像处理

(Digital Image Processing)

山东科技大学信息与电气工程学院

孙农亮





第4章 图像复原

第2讲 图像退化的数学模型及表达、常见退化模型



4.2 图像退化的数学模型

(4) 退化模型的空域表达

$$g(x,y) = f(x,y) * h(x,y) + n(x,y)$$

对数字化图像而言，计算卷积之前，需对各数据进行周期拓展。经过均匀取样和量化后，假设原数字图像 $f(x,y)$ 尺寸为 $A \times B$ ，点扩散函数 $h(x,y)$ 的大小为 $C \times D$ ，则都要拓展为大小为 $M \times N$ 的周期数据，其中 $M \geq A+C-1$ ， $N \geq B+D-1$ 。

$$f_e(x,y) = \begin{cases} f(x,y) & 0 \leq x \leq A-1 \text{ 和 } 0 \leq y \leq B-1 \\ 0 & A \leq x \leq M-1 \text{ 或 } B \leq y \leq N-1 \end{cases}$$

$$h_e(x,y) = \begin{cases} h(x,y) & 0 \leq x \leq C-1 \text{ 和 } 0 \leq y \leq D-1 \\ 0 & C \leq x \leq M-1 \text{ 或 } D \leq y \leq N-1 \end{cases}$$

$$n_e(x,y) = \begin{cases} n(x,y) & 0 \leq x \leq A-1 \text{ 和 } 0 \leq y \leq B-1 \\ 0 & A \leq x \leq M-1 \text{ 或 } B \leq y \leq N-1 \end{cases}$$

$$g_e(x,y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_e(m,n) h_e(x-m, y-n) + n_e(x,y)$$



4.2 图像退化的数学模型

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_{M-1} & \cdots & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_0 & \cdots & \mathbf{H}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{M-1} & \mathbf{H}_{M-2} & \cdots & \mathbf{H}_0 \end{bmatrix} \begin{bmatrix} f_e(0) \\ f_e(1) \\ \vdots \\ f_e(MN-1) \end{bmatrix} + \begin{bmatrix} n_e(0) \\ n_e(1) \\ \vdots \\ n_e(MN-1) \end{bmatrix}$$

由于 $h(x,y)$ 的周期性，使得 \mathbf{H} 成为一个块循环矩阵。
每一块如下所示：

$$\mathbf{H}_i = \begin{bmatrix} h_e(i,0) & h_e(i,N-1) & \cdots & h_e(i,1) \\ h_e(i,1) & h_e(i,0) & \cdots & h_e(i,2) \\ \vdots & \vdots & \ddots & \vdots \\ h_e(i,N-1) & h_e(i,N-2) & \cdots & h_e(i,0) \end{bmatrix}$$



4.2 图像退化的数学模型

例 4.1 请用一维离散信号为例，推导出矩阵 H 的表达式。假设一维离散信号 $f(x)$ 的采样点数是 3，系统点扩散函数 $h(x)$ 的采样点数为 2。

解答：为避免离散卷积的重叠，需进行周期拓展，最小周期可取 $M=3+2-1=4$ 。用 $f_e(x)$ 和 $h_e(x)$ 表示扩展之后的函数，将函数用零扩展补齐，即 $f_e(0)=f(0)$ ， $f_e(1)=f(1)$ ， $f_e(2)=f(2)$ ， $f_e(3)=0$ ； $h_e(0)=h(0)$ ， $h_e(1)=h(1)$ ， $h_e(2)=0$ ， $h_e(3)=0$ 。则它们的卷积为（ $x=0,1,2,3$ ）

$$g_e(x) = \sum_{m=0}^3 f_e(m) h_e(x-m)$$



4.2 图像退化的数学模型

$$g_e(x) = \sum_{m=0}^3 f_e(m) h_e(x-m)$$

$$= f_e(0) h_e(x-0) + f_e(1) h_e(x-1) + f_e(2) h_e(x-2) + f_e(3) h_e(x-3)$$

$$=[h_e(x-0), h_e(x-1), h_e(x-2), h_e(x-3)]$$

?	$f_e(0)$?
?	?	?
?	$f_e(1)$?
?	?	?
?	$f_e(2)$?
?	?	?
?	$f_e(3)$?

$$g_e(0) = [h_e(0-0), h_e(0-1), h_e(0-2), h_e(0-3)]$$

$$= [h_e(0), h_e(3), h_e(2), h_e(1)]$$

?	$f_e(0)$?
?	?	?
?	$f_e(1)$?
?	?	?
?	$f_e(2)$?
?	?	?
?	$f_e(3)$?



4.2 图像退化的数学模型

$$g_e(1) = [h_e(1-0), h_e(1-1), h_e(1-2), h_e(1-3)] \quad [h_e(1), h_e(0), h_e(3), h_e(2)]$$

$\begin{matrix} f_e(0) \\ ? \\ f_e(1) \\ ? \\ f_e(2) \\ ? \\ f_e(3) \end{matrix}$ $\begin{matrix} f_e(0) \\ ? \\ f_e(1) \\ ? \\ f_e(2) \\ ? \\ f_e(3) \end{matrix}$

$$g_e(2) = [h_e(2-0), h_e(2-1), h_e(2-2), h_e(2-3)] \quad [h_e(2), h_e(1), h_e(0), h_e(3)]$$

$\begin{matrix} f_e(0) \\ ? \\ f_e(1) \\ ? \\ f_e(2) \\ ? \\ f_e(3) \end{matrix}$ $\begin{matrix} f_e(0) \\ ? \\ f_e(1) \\ ? \\ f_e(2) \\ ? \\ f_e(3) \end{matrix}$

$$g_e(3) = [h_e(3-0), h_e(3-1), h_e(3-2), h_e(3-3)] \quad [h_e(3), h_e(2), h_e(1), h_e(0)]$$

$\begin{matrix} f_e(0) \\ ? \\ f_e(1) \\ ? \\ f_e(2) \\ ? \\ f_e(3) \end{matrix}$ $\begin{matrix} f_e(0) \\ ? \\ f_e(1) \\ ? \\ f_e(2) \\ ? \\ f_e(3) \end{matrix}$





4.2 图像退化的数学模型

$$\begin{array}{ccccccccccccccccc} \text{?} & \text{?} & \text{?} & h_e(0) & h_e(3) & h_e(2) & h_e(1) & \text{?} & \text{?} & h_e(0) & \text{?} & \text{?} & h_e(0) & 0 & 0 & h_e(1) & \text{?} & \text{?} \\ \text{?} & \text{?} & \text{?} & h_e(1) & h_e(0) & h_e(3) & h_e(2) & \text{?} & \text{?} & h_e(1) & \text{?} & \text{?} & h_e(1) & h_e(0) & 0 & 0 & \text{?} & \text{?} \\ \text{?} & \text{?} & \text{?} & h_e(2) & h_e(1) & h_e(0) & h_e(3) & \text{?} & \text{?} & h_e(2) & \text{?} & \text{?} & h_e(1) & h_e(0) & 0 & 0 & \text{?} & \text{?} \\ \text{?} & \text{?} & \text{?} & h_e(3) & h_e(2) & h_e(1) & h_e(0) & \text{?} & \text{?} & h_e(3) & \text{?} & \text{?} & h_e(1) & h_e(0) & 0 & 0 & \text{?} & \text{?} \end{array}$$

$$g=Hf$$



4.2 图像退化的数学模型

(5) 退化模型的频域表达

以上表达式是假定离散退化模型是线性移不变的，目的是在给定 $g(x,y)$ 和知道 $h(x,y)$ 、 $n(x,y)$ 的情况下，估计出原始图像 $f(x,y)$ 。但向量 \mathbf{g} 、 \mathbf{f} 、 \mathbf{n} 都是 MN 维的堆积向量，矩阵 \mathbf{H} 的大小是 $MN \times MN$ 。要解出原始图像，就要解 MN 个联立方程组，其计算量非常大，比如对 512×512 尺寸的图像而言， $MN \times MN = 262144 \times 262144$ ，即要直接求解 \mathbf{f} ，则要求解有 262144 个方程的联立方程组，这几乎是不可能的。

为此就要采取其他有效策略，方法之一是利用矩阵的特殊性质，将其对角化；更有效方法就是利用傅里叶变换卷积定理在频域中处理；实际上对角化的结果与傅里叶变换的结果是统一的。



4.2 图像退化的数学模型

$$g(x,y) = f(x,y) * h(x,y) + n(x,y)$$

$$G(u,v) = H(u,v)F(u,v) + N(u,v)$$

注意事项：



4.2 图像退化的数学模型

(6) 常见退化模型

1) 运动模糊退化模型

原因、类型：平移、旋转、匀速、变速等。这里讨论匀速直线运动。设 T 为快门时间， $x_0(t)$ 为位移的 x 分量。

$$g(x, y) = \boxed{\int_0^T f(x - x_0(t), y) dt}$$

对上式进行傅立叶变换【展开】，得

$$G(u, v) = \boxed{f(x, y)} e^{-j2\pi(ux+vy)} dx dy$$

$$= F(u, v) \boxed{\int_0^T} \exp\{-j2\pi(ux_0(t))\} dt = F(u, v) H(u, v)$$



4.2 图像退化的数学模型

$$H(u, v) = \int_0^T \exp\{-j2\pi(ux_0(t)\} dt$$

若 $x_0(t)=at/T$, 则

$$H(u, v) = \frac{T}{\pi ua} \sin(\pi ua) e^{-j\pi ua}$$

若 $x_0(t)=at/T$, $y_0(t)=bt/T$, 则

$$\begin{aligned} H(u, v) &= \int_0^T \exp\{-j2\pi [ux_0(t) + vy_0(t)]\} dt \\ &= \int_0^T \exp[-j2\pi(ua + vb)t / T] dt \\ &= \frac{T}{\pi(ua + vb)} \sin[\pi(ua + vb)] \exp[-j\pi(ua + vb)] \end{aligned}$$



4.2 图像退化的数学模型

2) 离焦模糊退化模型

(a) 圆柱模型 (基于几何光学)

$$h(x, y) = h(r) = \begin{cases} \frac{1}{\pi R^2} & r \leq R, \quad r^2 = x^2 + y^2 \\ 0 & \text{其他} \end{cases}$$

(b) 高斯模型 (基于波动光学)

$$h(x, y) = h(r) = \begin{cases} \frac{1}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}} & \\ 0 & \end{cases} = \begin{cases} \frac{1}{\pi R^2} e^{-\frac{r^2}{R^2}}, & r \leq R, r^2 = x^2 + y^2 \\ 0 & \end{cases}$$

(c) 优缺点



4.2 图像退化的数学模型

小结：几个概念，退化模型，线性移不变系统，空域表达，频域表达，常见模型

A photograph of a sunset over a body of water. The sky is a gradient from orange at the horizon to dark blue at the top. A single swan is swimming towards the right in the lower right foreground. In the upper left foreground, there is a bright, overexposed area that appears to be a lamp post or a light fixture.

謝

謝

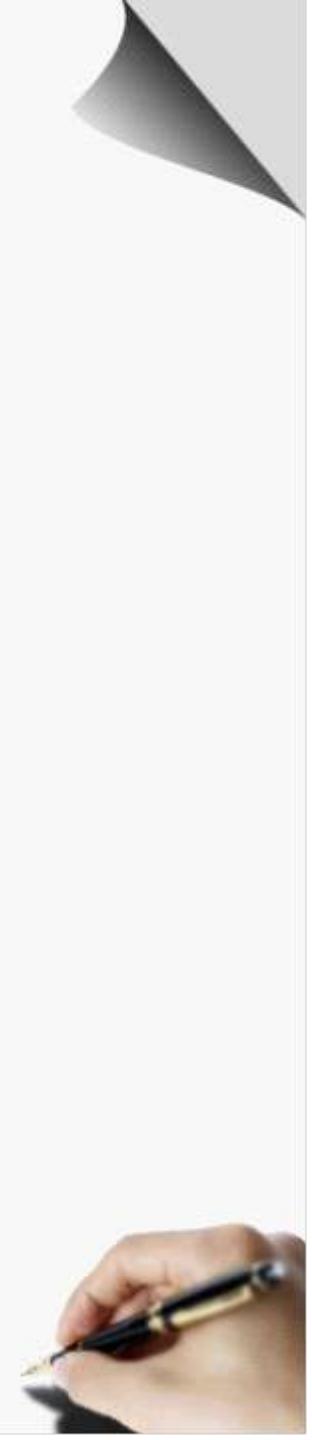


数字图像处理

(Digital Image Processing)

山东科技大学信息与电气工程学院

孙农亮





第4章 图像复原

第3讲 无约束复原 - 逆滤波方法



4.3 图像复原

图像退化过程可看成是一种变换， $g(x,y)=T[f(x,y)]$ ，由 $g(x,y)$ 求 $f(x,y)$ 就是求其逆变换的过程 $f(x,y)=T^{-1}[g(x,y)]$

在实际中， T^{-1} 有多种情况：

T^{-1} 不存在，即奇异

T^{-1} 存在，但不唯一

T^{-1} 存在，唯一，但 $g(x,y)$ 小的扰动就会引起 $f(x,y)$ 大的变动

T^{-1} 存在，唯一，但其解太复杂，或几乎不可解

T^{-1} 存在，唯一，无病态问题，且可求解



4.3 图像复原

(1) 无约束复原

当 T^{-1} 存在，唯一，无病态问题，原图像可精确求解；而当 T^{-1} 不存在，或存在但不可解时，原图像只能通过退化的 g 和对退化模型及噪声的某种了解或假设估计得到。这种估计是在某种最佳准则下的最佳估计，广义上分为无约束和有约束估计。

$$g = Hf + n \quad n = g - Hf$$

噪声是广义的，在没有先验知识的情况下，要找一个 f 的估计 \hat{f} ，在最小二乘方意义上使下式达到最小：

$$J(\hat{f}) = \|g - H \cdot \hat{f}\|^2$$



4.3 图像复原

$$\frac{\partial J(\hat{f})}{\partial \hat{f}} = -2H^T(g - H\hat{f}) = 0 \rightarrow H^T H \hat{f} = H^T g$$
$$\hat{f} = (H^T H)^{-1} H^T g$$

因为 H 是一方阵，并且设 H^{-1} 存在，则可求得 \hat{f} :

$$\hat{f} = H^{-1}(H^T)^{-1} H^T g = H^{-1}g$$

这种方法要求知道成象系统的表达式 H 。

根据前面所述， H 的尺寸很大，如 512×512 尺寸的图像， H 的尺寸为 262144×262144 ，对其求逆是不可解的，故要寻求合适的求解方法。

从 $G(u, v) = H(u, v) F(u, v) + N(u, v)$ 出发，若不考虑噪声，则上式可写成（逆滤波） $G(u, v) = H(u, v) F(u, v)$

$$\hat{F}(u, v) = G(u, v) / H(u, v)$$

$$\hat{f}(x, y) = F^{-1}[G(u, v) / H(u, v)]$$



4.3 图像复原

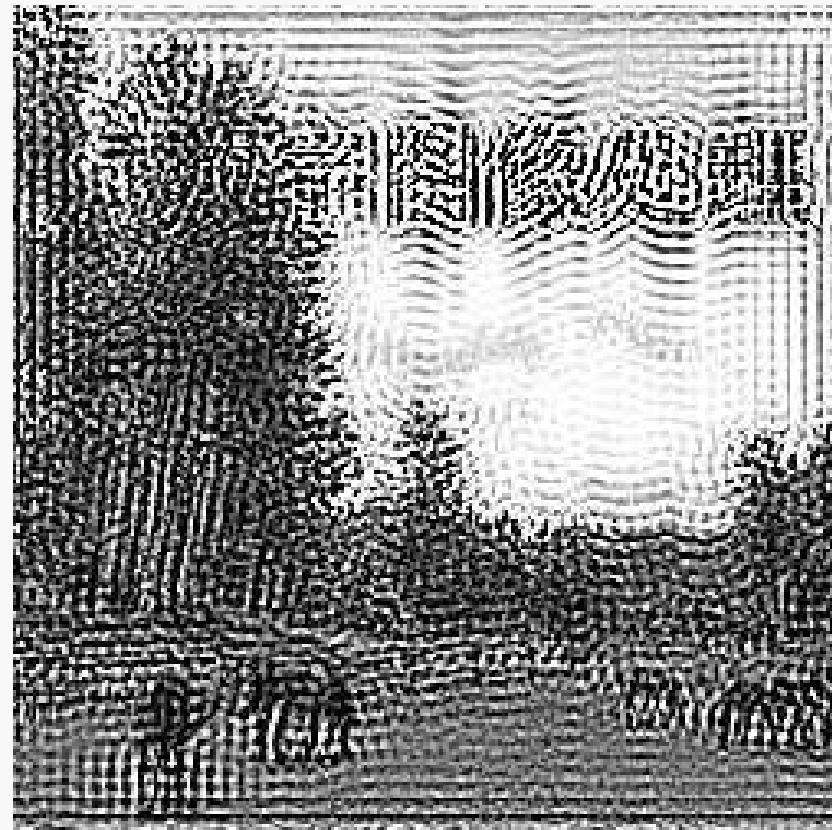
例 4.2 请用逆滤波法对已退化图像进行恢复，并分析结果
(假设系统响应是已知的)



退化图像



4.3 图像复原



用逆滤波直接回复的结果

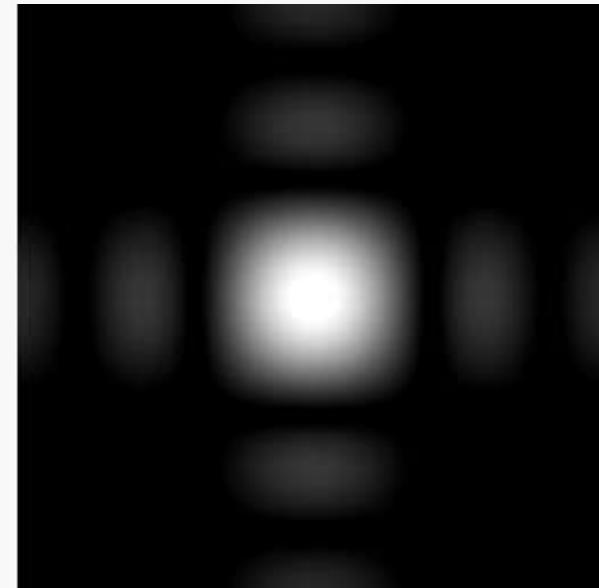


4.3 图像复原

分析 / 讨论

由逆滤波知：

- (a) $H(u, v)$ 在 UV 平面上取零或很小， $G(u, v) / H(u, v)$ 就会使恢复结果与预期的结果有很大差距；
- (b) 噪声带来更严重的问题（知道 H 也估计不准 f ）



$H(u, v)$ 的频谱图

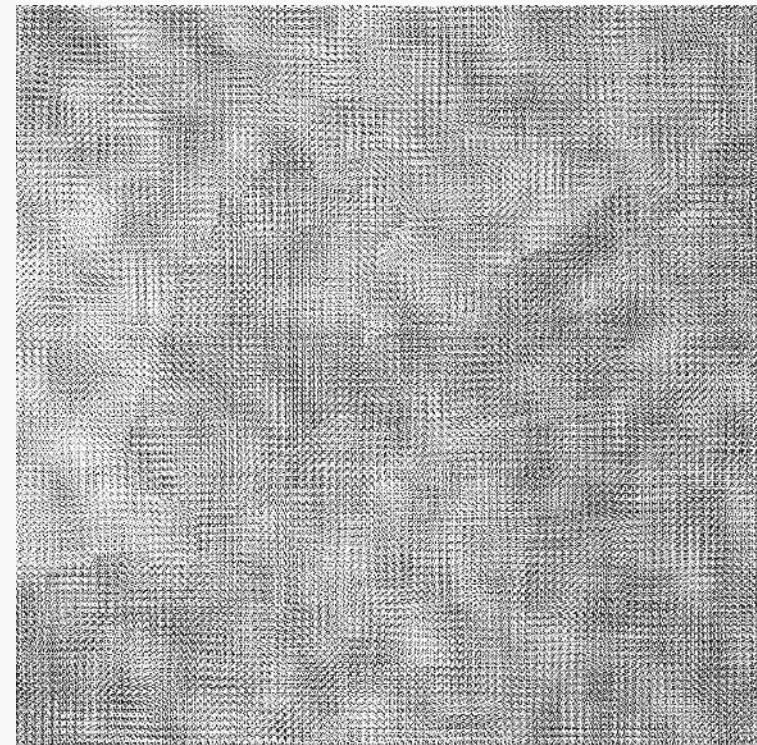
$$\hat{F}(u, v) = F(u, v) + \frac{N(u, v)}{H(u, v)} \quad u, v = 0, 1, \dots, M - 1$$

此式表明，无论怎样我们都无法得到原图像的精确值。不管噪声 $N(u, v)$ 是否已知，由于 $H(u, v)$ 很小的值和零点，都会使比值 $N(u, v) / H(u, v)$ 很大或无穷，即噪声掩盖了真实的图像值，使得恢复无意义。





4.3 图像复原



含高斯噪声的模糊图像逆滤波法直接恢复的效果



4.3 图像复原

改进：解决该问题的方法就是避开 $H(u,v)$ 的零点及小数值的点。

(a) 当 $H(u,v)$ 的值小于某个值 d 时取一个常数 k ，其他不变

$$M(u,v) = \begin{cases} k & \text{如 } H(u,v) \leq d \\ 1/H(u,v) & \text{其它} \end{cases}$$

(b) 在 $H(u,v)$ 的第一个零点之前取原值，之后取某个常数，如 1

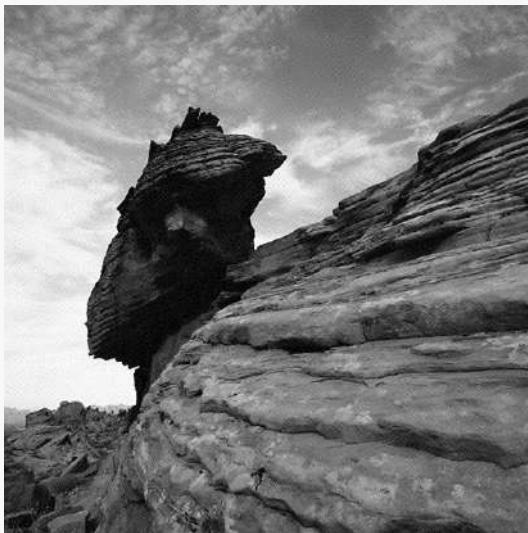
$$M(u,v) = \begin{cases} 1/H(u,v) & \text{如 } u^2 + v^2 \leq w_0^2 \\ 1 & \text{如 } u^2 + v^2 > w_0^2 \end{cases}$$



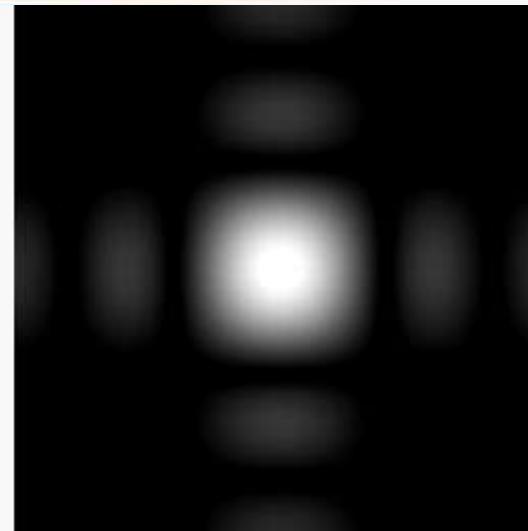
4.3 图像复原



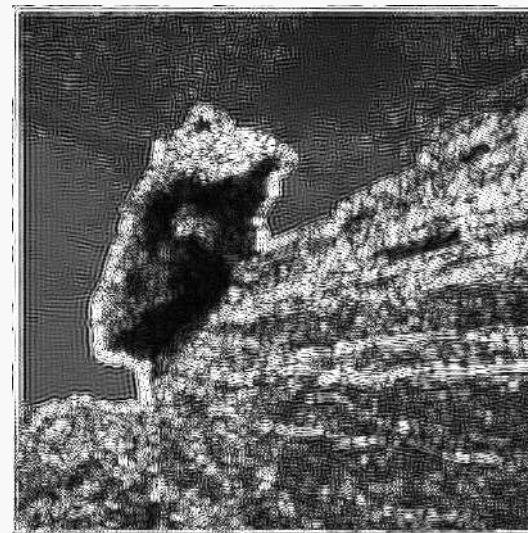
(a) 含有噪声的模糊图像



(c) 是将 $H(u,v)$ 中数值小于某个值或为零的点赋予某个常数 k ($k < 1$)



(b) 低通滤波器 $H(u,v)$ 的频谱图



采用逆滤波法恢复的图像

- (d) 是在频域原点附近取值 (第一个零点之前取原 $H(u,v)$ 值, 其余部分的值赋 1)

谢谢 谢谢

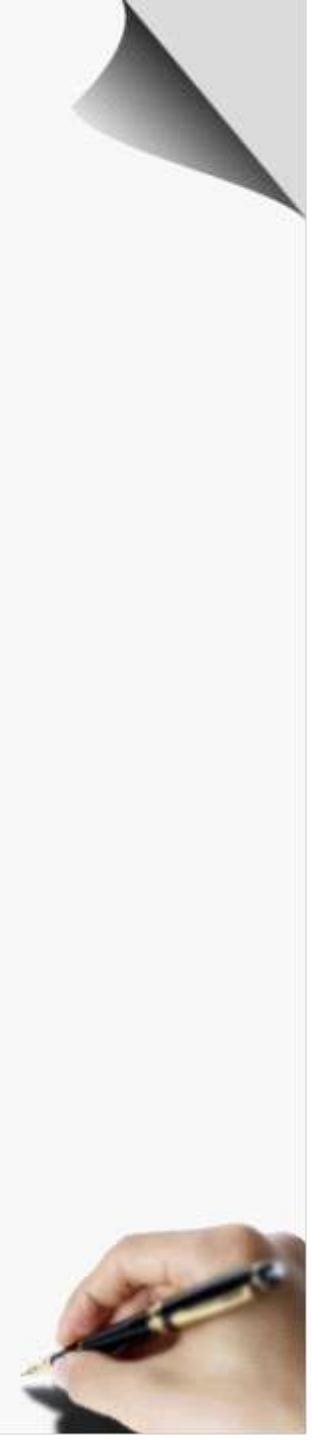


数字图像处理

(Digital Image Processing)

山东科技大学信息与电气工程学院

孙农亮





第4章 图像复原

第4讲 有约束复原 - 维纳滤波和约束最小平方滤波



4.3 图像复原

(2) 有约束恢复

在图像恢复过程中，为了在数学上更容易处理，常常附加某种约束条件，并在这些条件下使某个准则函数最小化。这类方法叫做有约束恢复，其中典型的方法有维纳滤波法和约束最小平方滤波法。

1) 维纳滤波法 (Wiener)，即最小均方误差滤波法

前提：图像和噪声都是随机过程，图像和噪声不相关。

基本思想：找到原图像 $f(x,y)$ 的一个估计值，使得估计值与原图像之间的均方误差在统计意义上最小。

$$e^2 = E\{[f(x,y) - \hat{f}(x,y)]^2\}$$



4.3 图像复原

$$P(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 + S_{nn}(u, v) / S_{ff}(u, v)}$$

图像估计值在频域的表达式为

$$\hat{F}(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 + S_{nn}(u, v) / S_{ff}(u, v)} G(u, v)$$

各项的含义，【不同的表示法】



4.3 图像复原

特点：对噪声有自动抑制作用

若噪声为白噪声，则 $S_{nn}(u,v)$ 近似为常数。当 $H(u,v)$ 在某处为零时，由于存在 $S_{nn}(u,v)/S_{ff}(u,v)$ ，不会出现被零除的情况，而由于分子含有项 $H^*(u,v)$ ，在任何 $H(u,v)=0$ 处，滤波器的增益恒等于 0。同时若在某一频谱区信噪比相当高或噪声为零时，维纳滤波器的效果接近逆滤波方法；而对信噪比很小的区域，滤波器的输出很小。这些都说明了维纳滤波避免了在逆滤波法中出现的对噪声的过多放大。在实际应用中， $S_{nn}(u,v)$ 、 $S_{ff}(u,v)$ 未知，特别是 $S_{ff}(u,v)$ ，解决的方法之一是令 $S_{nn}(u,v)/S_{ff}(u,v)$ 等于某个常数 k 。



4.3 图像复原

维纳滤波器进行图像恢复的主要步骤如下：

- (1) 假定系统的点扩散函数 $h(x,y)$ 是已知的，分别对退化图像 $g(x,y)$ 和 $h(x,y)$ 进行拓展，得 $g_e(x,y)$ 和 $h_e(x,y)$ ；
- (2) 分别计算 $g_e(x,y)$ 和 $h_e(x,y)$ 的傅里叶变换，得 $G_e(u,v)$ 和 $H_e(u,v)$ ；
- (3) 求 $H_e(u,v)$ 的复共轭和模 $|H_e(u,v)|^2$ ；
- (4) 选取 k 的初始值，($k \in [0.0001, 0.1]$)；
- (5) 根据上式计算 $F_e(u,v)$ ；
- (6) 对进行傅里叶反变换得原始图像的估计值 $f(x,y)$ ；
- (7) 选取另外一个 k 值，重复以上过程，并进行比较，直到找到最佳的 k 值以及对应的恢复图像。



4.3 图像复原

例 4.3 假设由匀速直线运动造成了图像模糊（仿真），图像中不含可视化噪声。试采用维纳滤波法和不同的参数，对图像进行复原。本例重点探讨点扩散函数的参数对恢复的影响。

解答：对相机与物体之间有相对匀速直线运动的情况，假设在一定的时间间隔里（相机的曝光时间）物体沿 x 、 y 方向移动的距离分别是 a 和 b 。当用 MATLAB7.0 编程实现时，所用参量为 θ 、 l 。它们与式 (4.2.22) 中的 a 、 b 的关系是 $\theta=\arctan(b/a)$, $l=(a^2+b^2)^{1/2}$, 当这两个参数确定后，点扩散函数的形式和参数也就定了。本例中取 $\theta=10^\circ$, $l=15 \text{ pixel}$ 。



4.3 图像复原

采用维纳滤波法进行恢复，因为没有加噪声，故 k 取了一个很小的值 $k=0.0001$ ($k=0$ 就变成了逆滤波)

其中一组是正确的值，即恢复时所用的参数与产生模糊图像时点扩散函数的参数值是一样的（此时 $\theta=10^\circ$ ， $l=15$ ），这是模拟了正确估计出退化的模型和相应的参数，恢复的图像如图 4.5(c) 所示。其他的几组参数要么角度值不正确，要么移动的距离不对，是对退化参数估计不准情形的模拟。图 4.5(d)、(e)、(f) 分别是 $\theta=10^\circ$ ， $l=10$ ； $\theta=10^\circ$ ， $l=20$ ； $\theta=5^\circ$ ， $l=15$ 时的恢复图像



4.3 图像复原



(a) 原图像



(b) $\theta = 10^\circ$, $l = 15$ 像素运动模糊图像



(c) $\theta = 10^\circ$, $l = 15$ 时恢复图像



(d) $\theta = 10^\circ$, $l = 10$ 时恢复图像



(e) $\theta = 10^\circ$, $l = 20$ 时恢复图像



(f) $\theta = 5^\circ$, $l = 15$ 时恢复图像





4.3 图像复原

例 4.4 采用维纳滤波法，对有噪声的离焦模糊图像复原。

解答：模拟离焦模糊的高斯模板（退化系统点扩散函数）的参数：方差为 5 pixel，模板是 7x7 pixel 的方模板。图像中含有加性高斯噪声，噪声的均值为零，方差为 0.001。

图 (a) 为原始图像，图 (b) 是含噪声的离焦模糊图像。

图 (c) 是用实际的噪声和原图像平均功率计算的功率比，即公式中的 $k=0.0029$ ，PSF 用真实值得到的恢复图像，图像较好；图 (d)、(e) 是用不准确的 k 值和不准确 σ 的值（分别为 $k=0.0003, \sigma=3$ 和 $k=0.03, \sigma=7$ ）恢复的情况，质量较差；图 (f) 是采用真实的 PSF，并用真实的噪信功率比恢复的结果，是所有恢复结果中最好的。



4.3 图像复原



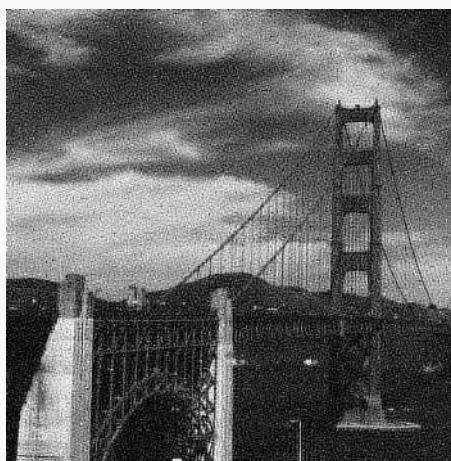
(a) 原始图像



(b) 含噪声离焦模糊图像



(c) PSF 用真实值, 用实际的噪声和原图像平均功率, 即 $k=0.0029, \sigma=5$



(d) 用偏小的值, $k=0.0003, \sigma=3$



(e) 用偏大的值, $k=0.03, \sigma=7$



(f) 用真实的 PSF 和真实的噪信功率比





4.3 图像复原

2) 约束最小平方滤波法

在恢复过程中，为减少噪声的影响，要建立某种基于平滑测度的最优准则。定义**准则函数 C** ，并在约束条件下使准则函数 C 最小

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

约束条件是：

$$\|g - H\hat{f}\|^2 = \|n\|^2$$



4.3 图像复原

拉氏算子采用以下形式

$$p(x, y) = \begin{matrix} ? & -1 & 0 & ? \\ ? & 1 & 4 & -1 & ? \\ ? & -1 & 0 & ? \end{matrix}$$

采用拉格朗日乘子法，并得到频域中以上问题的解为

$$\hat{F}(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 + s|P(u, v)|^2} G(u, v)$$

s 的求取和选择是关键，如何进行？



4.3 图像复原

迭代方式求解 s

定义一个中间变量 r ，即

$$r = g - H\hat{f}$$

相应的频域表达式为

$$R(u, v) = G(u, v) - H(u, v)\hat{F}(u, v)$$

$$\varphi(s) = \|r\|^2 = r^T r = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} r^2(x, y)$$

在给定精度因子 a 的情况下，若调整 s 使得下式成立，
则认为恢复达到了要求

$$\|n\|^2 - a \leq \|r\|^2 \leq \|n\|^2 + a$$



4.3 图像复原

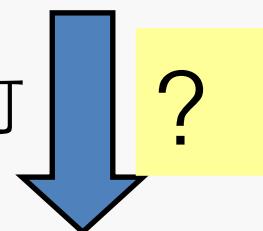
噪声的特性：

若对噪声图像具有先验知识，则可求其均值和方差

$$m_n = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} n(x, y)$$

$$\sigma_n^2 = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [n(x, y) - m_n]^2$$

实际上只要知道其均值和方差即可



$$\|n\|^2 = MN(m_n^2 + \sigma_n^2)$$



4.3 图像复原

约束最小平方滤波法恢复图像的 步骤：

(1) 根据先验知识，计算 $\|\mathbf{n}\|^2$ ；给 s 赋一初值；

$$(2) R(u,v) = \frac{s |P(u,v)|^2}{|H(u,v)|^2 + s |P(u,v)|^2} G(u,v)$$

(3) 对 $R(u,v)$ 求傅里叶反变换，得 $r(x,y)$ ，并计算 $\|r\|^2 = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} r^2(x,y)$

(4) 给定精度因子 a ，其选择须根据噪声情况确定， $a=(0.05 \sim 0.5) \|\mathbf{n}\|^2$ ；

(5) 若 $\|r\|^2 \leq \|\mathbf{n}\|^2 - a$ ，则增加 s ，若 $\|r\|^2 \geq \|\mathbf{n}\|^2 + a$ ，则减少 s ；并返回 $(\|\mathbf{n}\|^2 - a \leq \|r\|^2 \leq \|\mathbf{n}\|^2 + a)$

$$\hat{F}(u,v)$$

(7) 若 $\hat{F}(x,y)$ 成立，则停止迭代，并计算



4.3 图像复原

例 4.5 采用约束最小平方滤波器法，对例 4.4 中相同的退化图像进行恢复。

解答：模拟离焦模糊的高斯模板参数：方差为 5 pixel，模板是 7x7 pixel 的方模板。高斯噪声的均值为零，方差为 0.001。图 (a) 为原始图像，图 (b) 是含噪声的离焦模糊图像。根据公式计算，噪声的功率是 409.6。

图 (d) 是用偏小的噪声功率值和的值 ($\|n\|^2=205$ ， $\sigma=3$)，图 (e) 是用偏大的噪声功率值和的值 ($\|n\|^2=650$ ， $\sigma=7$) 分别恢复的情况； s 的初值取 0.001，其迭代时的步长取 0.01，精度因子 $a=0.1\|n\|^2$ 。

图 (f) 是采用真实的点扩散函数，和噪声功率 ($\|n\|^2=409.6$ ， $\sigma=5$) 时迭代求解的结果，精度因子 $a=0.1\|n\|^2$ ， s 的初值取 0.001，其迭代时的步长取 0.01，迭代终止时 $s=0.1201$ 。



4.3 图像复原



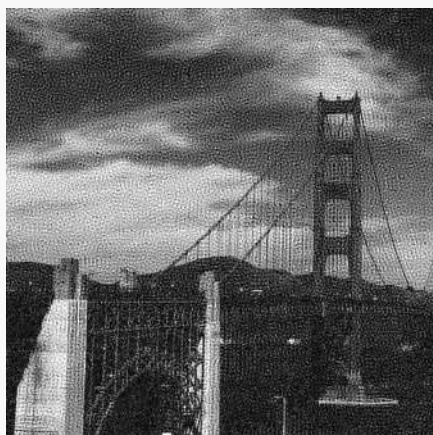
(a) 为原始图像



(b) 含噪声的离焦模糊图像



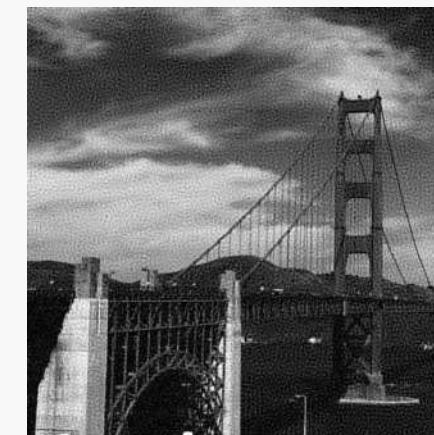
(c) $s=1$ ， $\sigma=5$ ，直接计算



(d) 用偏小的值 ($\|n\|^2=205$, $\sigma=3$)



(e) 用偏大的值 ($\|n\|^2=650$, $\sigma=7$)



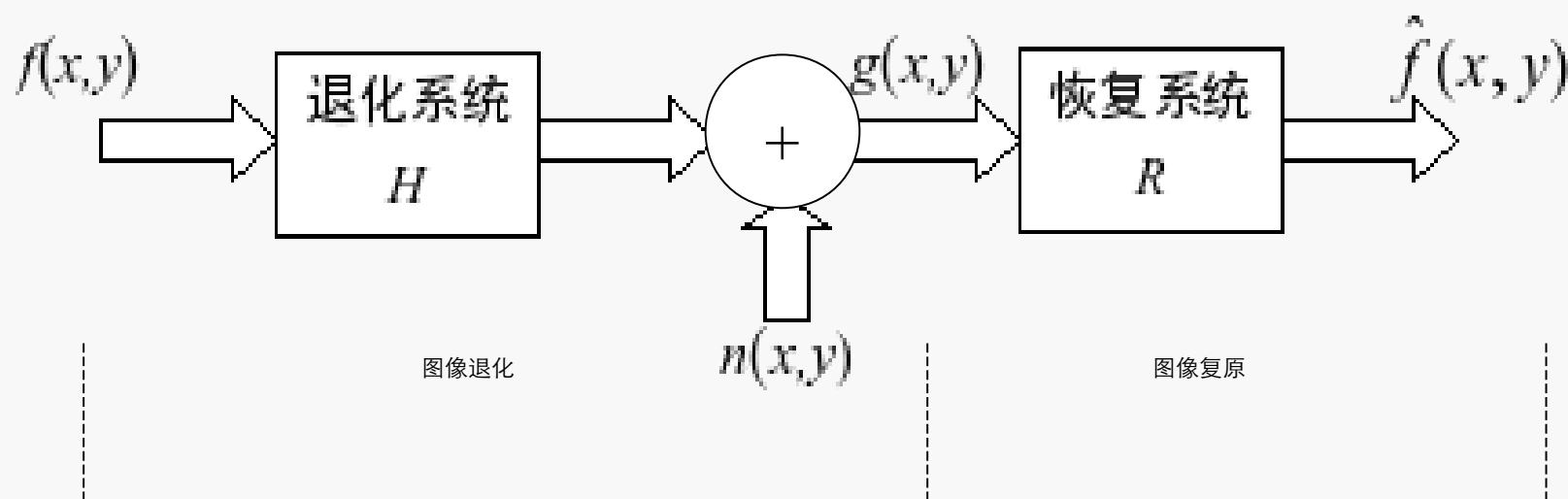
(f) 准确的值 ($\|n\|^2=409.6$, $\sigma=5$)





4.3 图像复原

本章小结



A photograph of a sunset over a body of water. In the foreground, a single black swan is swimming towards the right. On the left side of the frame, a street lamp is illuminated, casting a warm glow. The sky is filled with soft, orange and yellow hues.

謝 謝



图像压缩编码

山东科技大学
郑永果教授



第五章 图像压缩编码

讲解内容

1. 图像压缩的概念、目的和意义
2. 图像的霍夫曼编码、算术编码、行程编码、位平面编码、LZW编码、预测编码、变换编码等方法
3. 图像压缩的标准及发展现状

目的

1. 了解图像压缩的目的和意义，熟悉图像压缩评价方法；
2. 掌握常用图像编码方法
3. 了解图像压缩的标准及发展现状



第五章 图像编码与压缩

数据压缩研究内容包括数据的表示、传输、变换和编码方法，目的是减少存储数据所需的空间和传输所用的时间。

图像编码与压缩就是对图像数据按一定的规则进行变换和组合，达到以尽可能少的代码（符号）来表示尽可能多的图像信息。



数字图像的数据量往往非常大。以 1024×1024 的图像为例，8bit量化的灰度图像需要1MB的数据量，24bit量化的彩色图像需要3MB的数据量。而实际使用的图像不是单独存在的，往往是连续、多频谱的图像，这无疑给图像的存储、处理和传输带来极大困难。

在现代通信中，图像传输已成为重要内容之一。采用编码压缩技术，减少传输数据量，是提高通信速度的重要手段。

可见，没有图像编码与压缩技术的发展，大容量图像信息的存储与传输是难以实现的，多媒体、信息高速公路等新技术在实际中的应用会遇到很大困难。



5.1 图像压缩的基本原理

- 图像压缩所解决的主要问题是尽量减少表示数字图像时所需要的数据量。减少数据量的基本原理是去除其中多余的数据。
- “数据”与“信息”

“数据”又称“消息”，是传送信息的手段，而“信息”是数据中所包含的有意义的内容。

例：“踏花归来马蹄香”，七个字包含的信息也可以用这样一句话来描述：“我赏花回来了，因为花很香，所以连马蹄也变得香了起来”。

这就是包含了数据冗余



5.1.1 图像信息的冗余

从信息论观点看，描述图像信源的数据由有用数据和冗余数据两部分组成。

冗余数据有：编码冗余、像素间冗余、心理视觉冗余3种。

(1) 如果图像的灰度级在编码时用的编码符号数多于表示每个灰度级实际所需的符号数，则用这种编码得到的图像中包含编码冗余。

(2) 图像的各个像素不是孤立存在的，多数像素可以比较方便的由其邻近像素的值预测出来，每个独立的像素所携带的信息相对较少，换由单个像素对图像的视觉贡献有很多是冗余的。

(3) 在正常的视觉处理过程中各种信息的相对重要程度不同，那些对于人的视觉来说不十分重要的信息称为心理视觉冗余，



如果能减少或消除其中的1种或多种冗余，就能取得数据压缩的效果。因此图像信息的压缩是可能的。但到底能压缩多少，除了和图像本身存在的冗余度大小有关外，很大程度取决于对图像质量的要求。



- 广播电视 压缩比3：1
- 可视电话 压缩比1500：1
- 目前高效图像压缩编码技术已能用硬件实现实时处理，在广播电视、工业电视、电视会议、可视电话、传真和互连网等多方面得到应用。



5.1.2 图像编码压缩的分类

根据解压重建后的图像和原始图像之间是否具有误差，图像编码压缩分为无误差（亦称无失真、无损、信息保持）编码和有误差（有失真或有损）编码两大类。

根据编码作用域划分，图像编码为空间域编码和变换域编码两大类。





5.1.3 图像编码模型

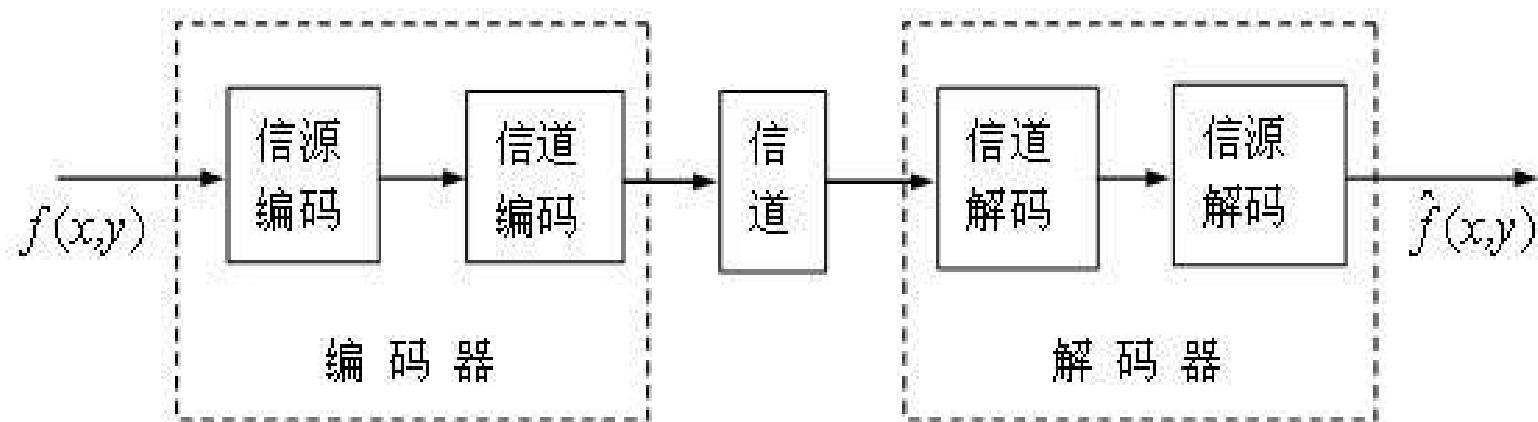


图5.2 图像压缩解压系统模型

一个图像压缩系统主要包括两个通过信道级连的结构模块，编码器和解码器，如图5.2所示

编码器对输入图像进行某种形式的压缩编码，生成数码率小于原始图像的适合于在信道中传输的一组符号。该组符号通过信道到达接收端，成为解码器的输入。解码器根据规定规则对该组符号解码，得到输出图像。如果系统是无损压缩且信道中没有噪声，则是的准确复制品；否则在重构图像中会出现某种程度的失真。



图5.2的编码器包括“信源编码”和“信道编码”两部分，相应的解码器包括“信道解码”与“信源解码”两部分。“信道解码”是“信道编码”的逆运算，“信源解码”完成的是“信源编码”的逆运算。

信源编码器用于减少或消除输入图像的三种冗余，实现数据压缩。

信道编码器实际上是差错控制编码器，用于给信源编码器的输出增加冗余信息，从而使其在信道中传输时具有更强的抗干扰能力。

如果传输信道是无噪声的，则信道编解码器可以略去。

由于信源编码器输出中几乎不包含冗余信息，因此传输时对噪声有很高的敏感性，很小的噪声都有可能导致大量的误码。

通常是通过信道编码器给其中插入一些预先规定好的有规律的冗余信息，在接收端通过验证所收到的信息是否满足预先设定好的规律而判断在传输过程中是否出错，有时还能对已出现的错误进行纠正，从而减少信道噪声对信号传输的影响。



5.1.4 图像保真度准则

在图像压缩编码中，解码图像与原始图像可能会有差异，因此需要评价压缩后图像的质量。

描述解码图像相对原始图像偏离程度的测度一般称为保真度。常用的保真度准则可分为两大类：客观保真度准则和主观保真度准则。



(1) 客观保真度准则

- 最常用的客观保真度准则是原图像和解码图像之间的均方根误差和均方根信噪比。
- 令 $f(x, y)$ 代表大小为 $M \times N$ 的原图像， $\hat{f}(x, y)$ 代表解压缩后得到的图像，对任意 x 和 y ， $f(x, y)$ 和 $\hat{f}(x, y)$ 之间的误差定义为：

$$e(x, y) = \hat{f}(x, y) - f(x, y)$$



则均方根误差 e_{rms} 为：

$$e_{rms} = \left\{ \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]^2 \right\}^{1/2}$$

如果将 $\hat{f}(x, y)$ 看做原始图像 $f(x, y)$ 和噪声信号 $e(x, y)$ 的和，那么解压图像的均方根信噪比 SNR_{rms} 为：

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



实际使用中，经常将 SNR_{max} 归一化并用分贝(db)表示。令

$$\bar{f} = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)$$

则有

$$SNR = 10 \cdot \log \left\{ \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - \bar{f}]^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]^2} \right\}$$



如果令 $f_{\max} = \max[f(x, y)]$, $x=0, 1, \dots, M-1$; $y=0, 1, \dots, N-1$ 。可得到峰值信噪比PSNR为：

$$PSNR = 10 \cdot \log \left\{ \frac{f_{\max}^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]^2} \right\}$$



(2) 主观保真度准则

对具有相同客观保真度的不同图像，人的视觉可能产生不同的视觉效果。这是因为客观保真度是一种统计平均意义下的度量准则，对于图像中的细节无法反映出来，而人的视觉能够觉察出来。这种情况下，用主观的方法来评价图像的质量更为合适。

很多解压图最终是供人观看的，一种常用的方法是让一组（不少于20人）观察者观察图像并给该图像评分，将他们对该图像的评分取平均，作为这幅图像的质量。



5.2 统计编码方法

根据信源的概率分布特性分配可变长码，使平均码长非常接近于熵。这种压缩编码称为统计编码。

常用的统计编码方法有：霍夫曼编码，行程编码，算术编码等



5.2.1 图像冗余度和编码效率

根据Shannon无干扰信息保持编码定理，若对原始图像数据的信息进行无失真图像编码，压缩后平均码长存在一个下限，这个下限是图像信息熵H。理论上最佳信息保持编码的平均码长可以无限接近图像信息熵H。但总是大于或等于图像的熵H。

$$H = - \sum_{i=0}^{L-1} p_i \log_2 p_i$$

平均码长 $\bar{B} = \sum_{i=0}^{L-1} \beta_i p_i$ β_i 是灰度值为*i*的编码长度

冗余度为 $r = \frac{\bar{B}}{H} - 1$

编码效率为

$$\eta = \frac{H}{\bar{B}} = \frac{1}{1+r}$$



5.2.2 霍夫曼编码

Huffman编码是1952年由Huffman提出的一种编码方法。这种编码方法是根据信源数据符号发生的概率进行编码的。

思想：在信源数据中出现概率越大的符号，编码以后相应的码长越短；出现概率越小的符号，其码长越长，从而达到用尽可能少的码符表示信源数据。它在无损变长编码方法中是最佳的。下面通过实例来说明这种编码方法。

设输入编码为 $X = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ ，其频率分布分别为 $P(x_1)=0.4$ ， $P(x_2)=0.3$ ， $P(x_3)=0.1$ ， $P(x_4)=0.1$ ， $P(x_5)=0.06$ ， $P(x_6)=0.04$ 。求其最佳霍夫曼编码

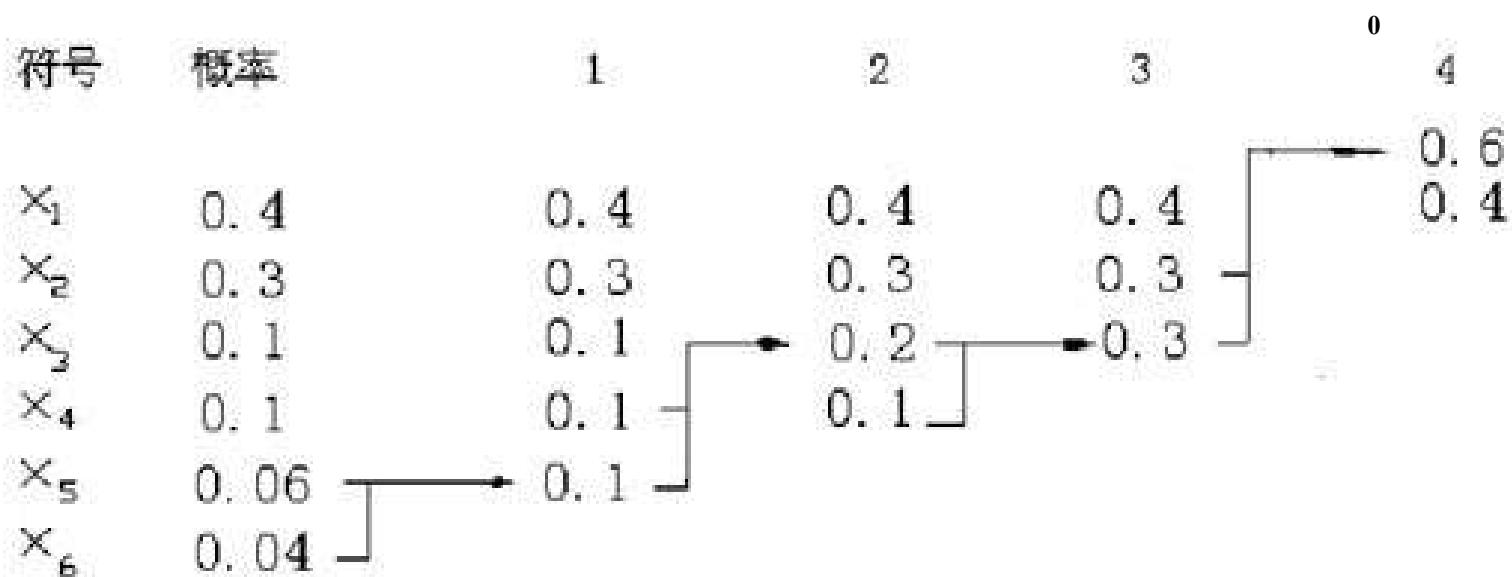
$$W = \{w_1, w_2, w_3, w_4, w_5, w_6\}$$



编码方法是：

- ①把输入符号按出现的概率从大到小排列起来，接着把概率最小的两个符号的概率求和；
- ②把它（概率之和）同其余符号概率由大到小排序，然后把两个最小概率求和；
- ③重复②，直到最后只剩下两个概率为止。

在上述工作完毕之后，从最后两个概率开始逐步向前进行编码。对于概率大的消息赋予0，小的赋予1。

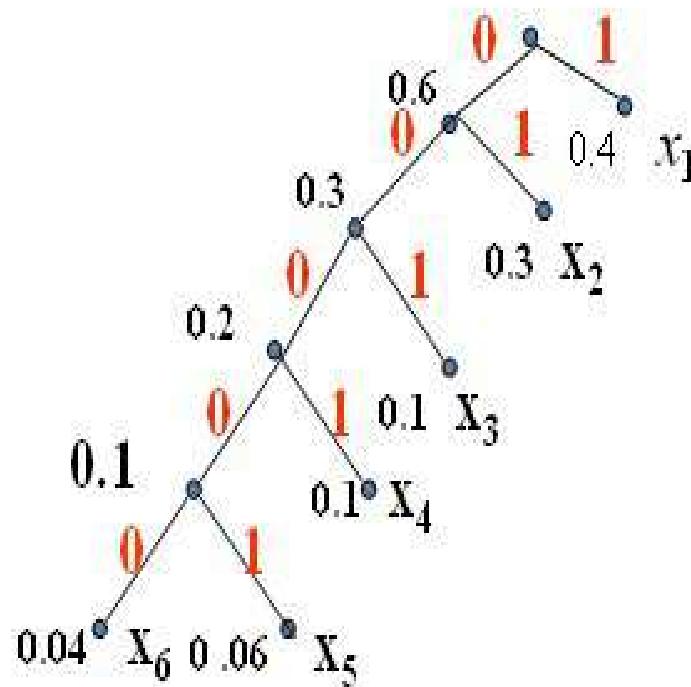


符号 x_i	x_1	x_2	x_3	x_4	x_5	x_6
概率 $P(x_i)$	0.4	0.3	0.1	0.1	0.06	0.04
编 码 w_i	1	00	011	0100	01010	01011



用二叉树方法实现 Huffman 编码方法也较为便

利。



课后题：计算该信源的熵、编码后的平均码长，

思考题：对于同一图像采用Huffman编码，编码是否唯一？



图像压缩编码

山东科技大学
郑永果教授



5.2.3 费诺—仙依编码

由于 **Huffman** 编码需要多次排序，当元素很多时十分不便，为此费诺和仙依分别单独提出类似的方法，使编码方法更简单。

具体编码方法如下：

(1) 把 $x_1 \sim x_n$ 按概率由大到小、由上到下排成一列，
然后把 $x_1 \sim x_n$ 分成两组 $x_1 \sim x_k, x_{k+1} \sim x_n$ ，并使得

$$\sum_{i=1}^k p(x_i) \approx \sum_{i=k+1}^n p(x_i)$$



- (2) 给两组中的 x_i 赋值，将概率大的一组赋为0，概率小的一组赋为1。
- (3) 把两组分别按(1)(2)分组、赋值，不断重复，直到每组只有一种输入元素为止。将每个 x_i 所赋的值依次排列起来就是费诺—仙依编码。



输入	概率					
x_1	0. 4	1				1
x_2	0. 3		0			00
x_3	0. 1			0	0	0100
x_4	0. 1	0	1		1	0101
x_5	0. 06			1	0	0110
x_6	0. 04				1	0111

费诺—仙依编码



5.2.4 算术编码

理论上，用霍夫曼方法对源数据流进行编码可达到最佳编码效果。但由于计算机中存储、处理的最小单位是位，因此在一些情况下，实际压缩比与理论压缩比的极限相去甚远。

例如：源数据流由X和Y两个符号构成，它们出现的概率分别是 $2/3$ 和 $1/3$ 。理论上字符X的熵确定的最优码长为：

$$H(X) = -\log_2(2/3) = 0.585 \text{ bit}$$

字符Y的最优码长为：

$$H(Y) = -\log_2(1/3) = 1.58 \text{ bit}$$



若要达到最佳编码效果，相应于字符X的码长为0.585位；字符Y的码长为1.58位。

用Huffman方法对这两个字符进行编码，得到X、Y的代码分别为0和1。显然，对于概率较大的字符X不能给予较短的代码。

这就是实际编码效果不能达到理论压缩比的原因所在。

算术编码把要压缩处理的整段数据映射到一段实数半开区间 $[0,1)$ 内的某一段，构造出小于1且大于或等于0的数值。这个数值是输入数据流的唯一可译码。



例：对一个5符号信源 $A=\{a_2a_3a_1a_3a_4\}$ ，各符号出现的概率和设定的取值范围如下：

字符	概率	范围
a_1	0.2	[0.0, 0.2)
a_2	0.2	[0.2, 0.4)
a_3	0.4	[0.4, 0.8)
a_4	0.2	[0.8, 1.0)

“范围”给出了符号的赋值区间。这个区间是根据符号发生的概率划分的。具体把 a_1 、 a_2 、 a_3 、 a_4 分配在哪个区间范围，对编码本身没有影响，只要保证编码器和解码器对符号的概率区间有相同的定义即可。



根据上述定义，若数据流的第一个符号为 a_2 ，由符号概率取值区间的定义可知，代码的实际取值范围在[0.2, 0.4)之间，亦即输入数据流的第一个符号决定了代码最高有效位取值的范围。然后继续对源数据流中的后续符号进行编码。每读入一个新的符号，输出。数值范围就将进一步缩小。读入第二个符号 a_3 取值范围在区间的[0.4, 0.8)内。但需要说明的是由于第一个符号 a_2 已将取值区间限制在[0.2, 0.4)的范围内，因此， a_3 的实际取值是在前符号范围[0.2, 0.4)的[0.4, 0.8)处。



将[0.2, 0.4)扩充为整个字符串的取值区间，第二个字符占据的取值空间是[0.2, 0.4)范围内的[0.4, 0.8)部分，即
[0.28, 0.36)。 $0.28=0.2+(0.4-0.2)*0.4$; $0.36=0.2+(0.4-0.2)*0.8$;

第三个字符a1是在前两个字符a2a3的基础上得到的，因此，它应按a1的概率对[0.28, 0.36)进行再一步细分，即
[0.28, 0.36)的[0.0, 0.2)部分，得到新的区间[0.28, 0.296)。
 $0.28=0.28+(0.36-0.28)*0.0$; $0.296=0.28+(0.36-0.28)*0.2$;

同理可得到第四个字符a3的取值空间[0.2864, 0.2928); 最后一个字符a4的取值空间为[0.29152, 0.2928)。



总结计算规律：

$$N_s = F_s + C_l * L \quad (5.1)$$

$$N_e = F_s + C_r * L \quad (5.2)$$

式中， N_s 为新子区间的起始位置； F_s 为前子区间的起始位置； C_l 为当前符号的区间左端； N_e 为新子区间的结束位置； C_r 为当前符号的区间右端； L 为前子区间的长度。



每输入一个符号，都将按事先对概率范围的定义，在逐步缩小的当前取值区间上，按式(5.1)和(5.2)式确定新的范围上、下限。重复上述编码过程，直到输入数据流结束。

字符	区间长度	范围
a_2	0.2	[0.2, 0.4)
a_3	0.08	[0.28, 0.36)
a_1	0.016	[0.28, 0.296)
a_3	0.0064	[0.2864, 0.2928)
a_4	0.00128	[0.2915, 0.2928]



编码序列



a_2

a_3

a_1

a_3

a_4

1

0.4

0.36

0.296

0.2928

a_4

a_4

a_4

a_4

a_4

a_3

a_3

a_3

a_3

a_3

a_2

a_2

a_2

a_2

a_2

a_1

a_1

a_1

a_1

a_1

0

0.2

0.28

0.28

0.2864





由此可见，随着符号的输入，代码的取值范围越来越小。当字符串 $A=\{a_2a_3a_1a_3a_4\}$ 被全部编码后，其范围在[0.2915, 0.2928]内。换句话说，在此范围内的数据代码都唯一对应于字符串“ $a_2a_3a_1a_3a_4$ ”。我们可取这个区间的下限 0.2915 作为对源数据流“ $a_2a_3a_1a_3a_4$ ”进行压缩编码后的输出代码，这样，就可以用一个浮点数表示一个字符串，达到减少所需存储空间的目的。



若为减少码长，可用0.292表示，这样三位十进制数字表示五个符号的消息，其平均码长为0.6个十进制数字/符号，而信源熵为0.58个十进制数字/符号，十分接近无噪声信源编码定理所设定的理论极限。

编码序列的长度越长，算术编码就越接近无噪声编码定理所设定的理论极限。但是在实际上有两个因素使编码效率无法达到这个极限值：①为了将各个消息分开，必须增加消息结束指示符；②使用的算法精度是有限的。



按这种编码方案得到的代码，其解码过程的实现比较简单。根据编码时所使用的字符概率区间分配表和压缩后的数值代码所在的范围，可以很容易地确定代码所对应的第一个字符。在完成对第一个符号的解码后，设法去掉第一个符号对区间的影响，再使用相同的方法找到下一个符号。重复以上的操作，直到完成解码过程。



假设以0.292作为码字，则其解码过程如下：

- (1) 根据码字所在范围确定消息序列的第一个码字：
0.292位于[0.2, 0.4)之间，因此第一个码字为a2。
- (2) 消除已译码字的影响，确定下一个码字。

需要将 第一个码字的影从码字0.292中消除。其运算为编码的逆运算，将a2的取值下限0.2从码字0.292中减掉，再除以a2的取值宽度
 $0.2(=0.4-0.2)$ ，得到新的码字0.46。

- (3) 重复步骤(1)与(2)的做法，直到码处理完成为止。



本章重点：1. 图像压缩的概念，图像评价的方法

2. 图像的霍夫曼编码、算术编码、行程编码、
预测编码、变换编码等方法

3. 图像压缩的标准

作业：第五章1,2,3,4



数字图像处理

(Digital Image Processing)

山东科技大学



图像压缩编码

山东科技大学
郑永果教授



第五章 图像压缩编码

已讲解内容

1. 图像压缩的概念

图像压缩，“数据”与“信息”

数据冗余，编码冗余、像素间冗余、心理视觉冗余

无误差（亦称无失真、无损、信息保持）编码和有误差
(有失真或有损) 编码

图像编码模型，信源编码，信源解码，信道编码，信道
解码

图像保真度准则，客观保真度准则，主观保真度准则

2. 统计编码方法。图像的霍夫曼编码、算术编码、

本次课内容：无误差压缩编码，有误差压缩编码

行程编码、位平面编码、预测编码、变换编码等方法



5.3 无误差压缩编码

根据在压缩过程中是否有信息的损失可将图像压缩编码分为无误差和有误差编码两类。无误差编码在解码时能够完全恢复原始图像信息，但通常压缩比有限；有误差编码不能完全恢复原始图像信息，对信息来讲是有损失的，但这种损失通常是人们所不易察觉的，通常有误差编码的压缩比较高。

无误差编码可分为变长码和定长码。

定长码采用相同的位数(bit)对数据进行编码，大多数存储数字信息的编码系统都采用定长码，最常用的有行程编码和LZW编码。

变长码是基于统计方法进行的编码。根据图像中各像素出现的概率不同，采用不同位数对像素进行编码，对出现概率低的像素编码位数多，出现概率高的像素编码位数少，从而达到数据压缩的目的，最常用的是哈夫曼编码和算术编码。采用变长码能够消除编码冗余。



5.3.1 位平面编码

位平面编码：

(1) 将多灰度值图像分解成一系列二值图。

(2) 对每一幅二值图用二元压缩方法进行压缩。

这种技术除能消除或减少编码冗余外也能消除或减少图像中的像素间冗余。



1、位平面分解

对一幅用多个比特表示其灰度值的图像来说，其中的每个比特可看做表示了一个二值的平面，也称位面。例如，一幅其灰度级用8比特表示的图像有8个位面，一般用位面0代表最低位面，位面7代表最高位面。

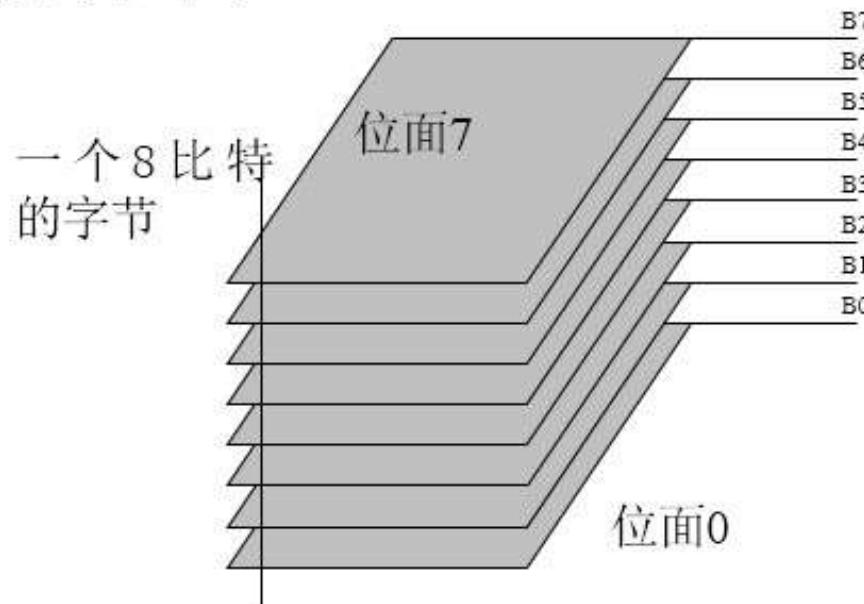


图5.6 图像的位面表示和位面图



对一幅8比特灰度级的图像，当代表一个像素灰度值字节的最高位比特为1时，该像素的灰度值必定大于或等于128，当最高位比特为0时，该像素的灰度值必定小于128，所以最高位面（位面7）相当于把原图灰度值分成0–127和128–255两个范围。

同理，位面6相当于把原图灰度值分成0–63及128–191和64–127及192–255这4个范围，并将前两者标为黑，后两个标为白，其他各图可依次类推。

由此可见，越是高的位面越包含了宏观的视觉可见的有意义信息，越是低的位面越多包含局部的细节，很多情况下也常认为是噪声。



m比特灰度级的图像中像素的灰度值可用以下多项式表示：

$$a_{m-1}2^{m-1} + a_{m-2}2^{m-2} + \dots + a_12^1 + a_02^0$$

(5.3.1)

根据式5.3.1，把一幅灰度图分解成一系列二值图集合的一种简单方法就是把上述多项式的m个系数分到m个1比特的位平面中去，如将每个像素的第*i*个比特集合在一起就能得到图像的第*i*个位面。



2、位平面的编码

位平面分解后得到的是二值图，即像素值只有0或1两种，并且许多位面图中有大片的连通区域，另外，像素值为0或1的两种像素在平面上是互补的。这些特性都可用来帮助进行编码。例如常数块编码、行程编码等。



常数块编码是利用专门的码字表达全是0或1的连通区域，以达到压缩的目的。它将图像分成全黑、全白或混的 $m \times n$ 尺寸的块。出现频率最高的类赋予**1**比特码字**0**，其他两类赋予**2**比特码字**10**和**11**（即哈夫曼编码）。由于原来需要 $m \times n$ 比特表示的常数块现在只用1比特或2比特码字表示，就达到了压缩的目的。当然这里赋给混合块的码只是前缀码，后面还需要跟上 $m \times n$ 比特表示的模式。



行程编码

另一种对常数区编码的有效方法是用一系列描述黑或白像素行程的长度来表示二值图或位平面的每一行，这种称为行程编码。它的基本思想是对一组从左向右扫描得到的连续的0或1行程用它们的长度来编码。这里需要建立确定行程值的协定，常用的方法为指出每行第一个行程的值，或者设每行都有白色行程开始。



5.3.2 二值图像编码

二值图像就是只有黑白两种灰度级特殊灰度图像，例如文件、地图、报纸等等。根据二值图像结构简单，黑、白像素区域多为连续分布等特征，可对二值图像进行压缩。

如果对每个像素用一位二进制码0或1（白像素为1，黑像素为0）表示，则称为直接编码，这是表示一幅二进制图像的比特数等于该图像的像素数。例如一幅 1024×768 的图像需要的存储空间为 1024×768 比特（98304字节即96kB），直接编码对数据量是没有压缩效果的，通常把直接编码得到的比特数称为原始数据大小。

常用的二值图像编码主要有跳白块编码、方块编码、识别编码等。



1、跳白块编码

基本跳白块编码是利用二值图像含有大量白色区域这一特点提出的编码方法。

将图像每行分成若干子块，每块包含N个像素。如果某块全部是白色像素，则该块用1比特字“0”表示；否则，如果某块至少含有一个黑色像素，则该块用N+1比特表示：前缀码“1”加该块的直接编码。

例如：原始数据为000000011100000，1表示该像素颜色为黑色，0为白色。设N=5，则编码后为：
01001110



2、方块编码

方块编码就是把整个图像分成等大小的子块，然后按每块内像素的不同排列所出现的概率分配不桶长度的字码，概率大的分配短字码，概率小的分配长码，从而使得平均码长达到最短（即哈夫曼编码）。由于和跳白块编码相类似，方块编码有时也被称作二维跳白块编码。



3、识别编码

早在1974年，Ascher和Nagy提出了利用模式匹配技术进行二值图像压缩编码的概念。对于打印机和印刷的文字来说，可以利用图像识别的方法来大大提高压缩比。如果发送端能够识别字符，则可以为每个字符分配一个定长的二进制码字，传送时只传送该码字即可，在接收端将码字恢复成原来的字符就完成了解码。



5.3.3 行程编码

行程编码又称 RLE 压缩方法，其中 RLE 是 **Run-Length-Encoding** 的缩写。这种压缩方法广泛地应用于各种图像格式的数据压缩处理中，是最简单的压缩图像的方法之一。



(1) 概念

行程: 具有相同灰度值的像素序列。

(2) 编码思想

去除像素冗余，用行程的灰度和行程的长度代替行程本身。

例: 设重复次数为 iC , 重复像素值为 iP

编码为: $iCiP \ iCiP \ iCiP$

编码前: aaaaaaabbbbbcccccddd $\text{d}\text{e}\text{d}\text{d}\text{d}\text{a}$

编码后: 7a6b8c**3d1e3d2a**

分析: 对于有大面积色块的图像, 压缩效果很好; 对于纷杂的图像, 压缩效果不好, 最坏情况下, 会加倍图像数据。



5.3.4 无损预测编码

预测编码是通过去除图像像素间相关性来进行数据压缩的。

活动图像是一种三维数据源，在空间上有行与列两维，在时间上是一帧一帧出现的。比如电视图像，采用每秒25帧的更换速率，利用人眼的视觉暂留现象产生连续感。

静态图像，像素之间存在着相关性。对于活动图像，由于帧速很高，图像是连续的，连续的图像帧之间也存在着相关性。

帧内预测编码：利用同一幅图像中像素间相关性进行预测编码。

帧间预测编码：利用图像帧之间的相关性进行预测编码称为。也可以将帧内预测编码与帧间预测编码联合使用以提高压缩比。



➤ 一幅二维静止图像，设空间坐标(i, j)像素点的实际灰度为 $f(i, j)$ ， $\hat{f}(i, j)$ 是根据以前已出现的像素点的灰度对该点的预测灰度，也称预测值或估计值，计算预测值的像素，可以是同一扫描行的前几个像素，或者是前几行上的像素，甚至是前几帧的邻近像素。实际值和预测值之间的差值，以下式表示：

$$e(i, j) = f(i, j) - \hat{f}(i, j) \quad (5-3)$$



因为图像像素的灰度是连续变化的，所以在一个区域中，相邻像素之间灰度值的差别可能很小。如果我们只记录第一个像素的灰度，其他像素的灰度都用它与前一个像素灰度之差来表示，就能起到压缩的目的。

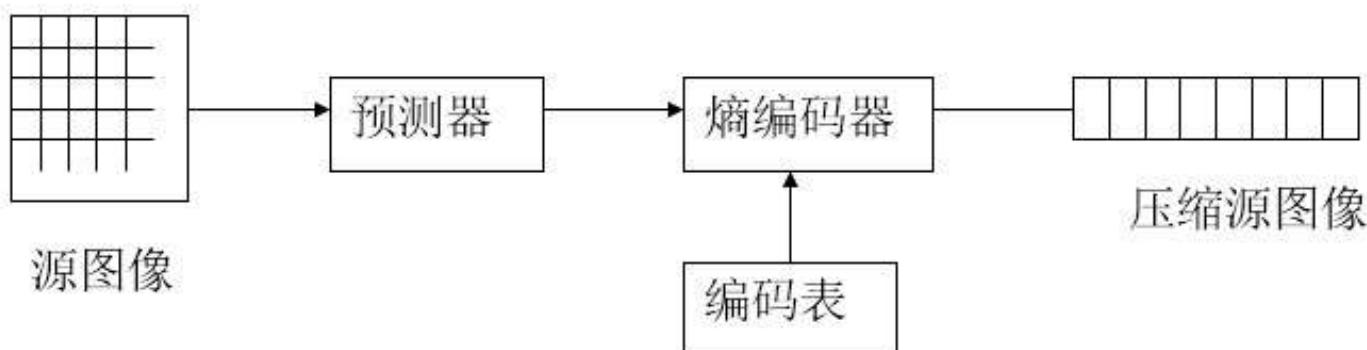
比如，**238, 2, 0, 1, 1, 3**，实际上这**6**个像素的灰度是**238, 240, 240, 241, 242, 245**。表示**238**需要**8**个比特，而表示差值只需**2**个比特，这样就实现了压缩。



- 由图像的统计特性可知，相邻像素之间有着较强的相关性。因此，其像素的值可根据以前已知的几个像素来估计，即预测。
- 预测编码是根据某一模型，利用以往的样本值对于新样本值进行预测，然后将样本的实际值与其预测值相减得到一个误差值，对于这一误差值进行编码。
- 如果模型足够好且样本序列在时间上相关性较强，那么误差信号的幅度将远远小于原始信号。
- 对差值信号不进行量化而直接编码就称之为无损预测编码。



➤ 无损预测编码器的工作原理图 如下：





► $\hat{f}(i, j)$ 由先前三点预测可以定义为：

$$\hat{f}(i, j) = a_1 f(i, j-1) + a_2 f(i-1, j-1) + a_3 f(i-1, j) \quad (5-4)$$

- 其中 a_1, a_2, a_3 称预测系数，都是待定参数。如果预测器中预测系数是固定不变的常数，称之为线性预测。
- 预测误差：

$$\begin{aligned} e(i, j) &= \hat{f}(i, j) - f(i, j) \\ &= f(i, j) - [a_1 f(i, j-1) + a_2 f(i-1, j-1) + a_3 f(i-1, j)] \end{aligned} \quad (5-5)$$



► 设 $a = f(i, j-1)$, $b = f(i-1, j)$, $c = f(i-1, j-1)$, $\hat{f}(i, j)$ 的预测方法如右图所示，可有8种选择方法：

	c	b	
	a	x	

选择方法	预测值 $\hat{f}(x, y)$
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	$(a+b)/2$



例：设有一幅图像， $f(i-1,j-1), f(i-1,j), f(i,j-1), f(i,j)$ 的灰度值分别为253,252,253,255，用上图第四种选择方法预测 $f(i,j)$ 的灰度值，并计算预测误差。

解： $\hat{f}(i,j) = a+b-c = f(i,j-1)+f(i-1,j)-f(i-1,j-1)$
 $= 253+252-253=252$

预测误差 $e(i,j) = f(i,j) - \hat{f}(i,j) = 255-252=3$

显然，预测误差 $e(i,j)=3$ 比像素的实际值 $f(i,j)=255$ 小的多，对3进行编码比对255直接编码将占用更少的比特位。



线性预测编码总结

在排序后的图像序列 $\{x_i\}(i=1, 2, \dots, N-1)$ 中，根据 x_1, x_2, \dots, x_{N-1} 对 x_N 作预测。令 x_N 的预测估计值为 x'_N ，假如 x'_N 是 x_1, x_2, \dots, x_{N-1} 的线性组合，则称对 x_N 的预测为线性预测。

假定二维图像信号 $x(t)$ 是一个均值为零、方差 σ^2 为的平稳随机过程， $x(t)$ 在 t_1, t_2, \dots, t_{N-1} 时刻的抽样集合为： $\{x_i\}=\{x_1, x_2, \dots, x_{N-1}\}$



则 t_N 时刻抽样值的线性预测为：

$$x'_N = \sum_{i=1}^{N-1} a_i x_i = a_1 x_1 + a_2 x_2 + \dots + a_{N-1} x_{N-1} \quad (5-6)$$

式中 a_i 为预测系数。

采用均方误差最小的准则，可得到 a_i 的最佳估计。设 x_N 的均方误差为：

$$E[e_N^2] = E[(x_N - x'_N)^2] = E\left\{[x_N - (a_1 x_1 + a_2 x_2 + \dots + a_{N-1} x_{N-1})]^2\right\} \quad (5-7)$$



为使 $E[e_N^2]$ 最小，上式两边对 a_i 求微分

$$\begin{aligned}\frac{\partial}{\partial a_i} E[e_N^2] &= \frac{\partial}{\partial a_i} E \left\{ [x_N - (a_1 x_1 + a_2 x_2 + \dots + a_{N-1} x_{N-1})]^2 \right\} \\ &= -2 E \left\{ [x_N - (a_1 x_1 + a_2 x_2 + \dots + a_{N-1} x_{N-1})] x_i \right\} \\ i &= 1, 2, \dots, N-1\end{aligned}$$

令 $\frac{\partial}{\partial a_i} E[e_N^2] = 0$

则 $E \left\{ [x_N - (a_1 x_1 + a_2 x_2 + \dots + a_{N-1} x_{N-1})] x_i \right\} = 0$

假设 x_i 和 x_j 的协方差为： $R_{ij} = E[x_i x_j]$ $i, j = 1, 2, \dots, N-1$

则(6.4.7)式可表示为： $R_{Ni} = a_1 R_{1i} + a_2 R_{2i} + \dots + a_{N-1} R_{(N-1)i}$



若所有的协方差 R_{ij} 已知，则可利用递推算法求解 $(N-1)$ 个预测系数 a_i 。

在线性预测编码中，若只采用 x_{N-1} 对 x_N 进行预测，称为前值预测。若采用同一行中 x_N 前的若干抽样值来对 x_N 预测，称为一维预测。若采用几行内的抽样值来预测 x_N ，称为二维预测。

根据前面的定义，假定图像是平稳的随机过程，因此可以用自相关系数代替 R_{ij} ，以求得预测系数。



非线性预测编码

线性预测编码的基础是假设图像全域为平稳的随机过程，自相关系数与像素在域中的位置无关。实际上，图像的灰度起伏始终是存在的，被描述像素和周围像素之间，含有多种多样的关系。线性预测系数 a_i 是一种近似条件下的常数，存在以下的缺点：

- (1) 对灰度有突变的地方，会有较大的预测误差，致使重建图像的边缘模糊，分辨率降低。



(2) 对灰度变化缓慢区域，其差值信号接近零，但因其预测值偏大而使重构图像有颗粒噪声。

为了改善图像质量，克服上述预测编码带来的缺点，非线性预测充分考虑了图像的统计特性和个别变化，尽量使预测系数与图像所处的局部特性相匹配，即预测系数随预测环境而变，所以称为自适应预测编码。



将式(5-6)改写成

$$x'_N = k \sum_{i=1}^{N-1} a_i x_i = k(a_1 x_1 + a_2 x_2 + \dots + a_{N-1} x_{N-1}) \quad (5-9)$$

这里 k 为自适应系数。一般情况下，令 $k=1$ ，但对灰度变化大的局部，由于预测值偏小，这时可令 $k=1.125$ ，以避免局部边缘被平滑；对灰度变化缓慢区，预测值可能偏大，这时可令 $k=0.875$ ，以消除颗粒噪声的影响。



数字图像处理

(Digital Image Processing)

山东科技大学



图像压缩编码

山东科技大学
郑永果教授



5.4 有误差压缩编码

无误差压缩编码通过消除数据间的编码冗余来达到编码压缩的目的，并且通过解码能够全恢复原始图像信息，但通常压缩比有限；

有误差编码，虽然有误差编码不能完全恢复原始图像信息，对信息来讲是有损失的，但这种损失通常是人们所不易察觉的。

有损编码是以丢失部分信息为代价来换取高压缩比。

常见的有误差压缩编码：预测编码和变换编码。本节将介绍有损预测编码。

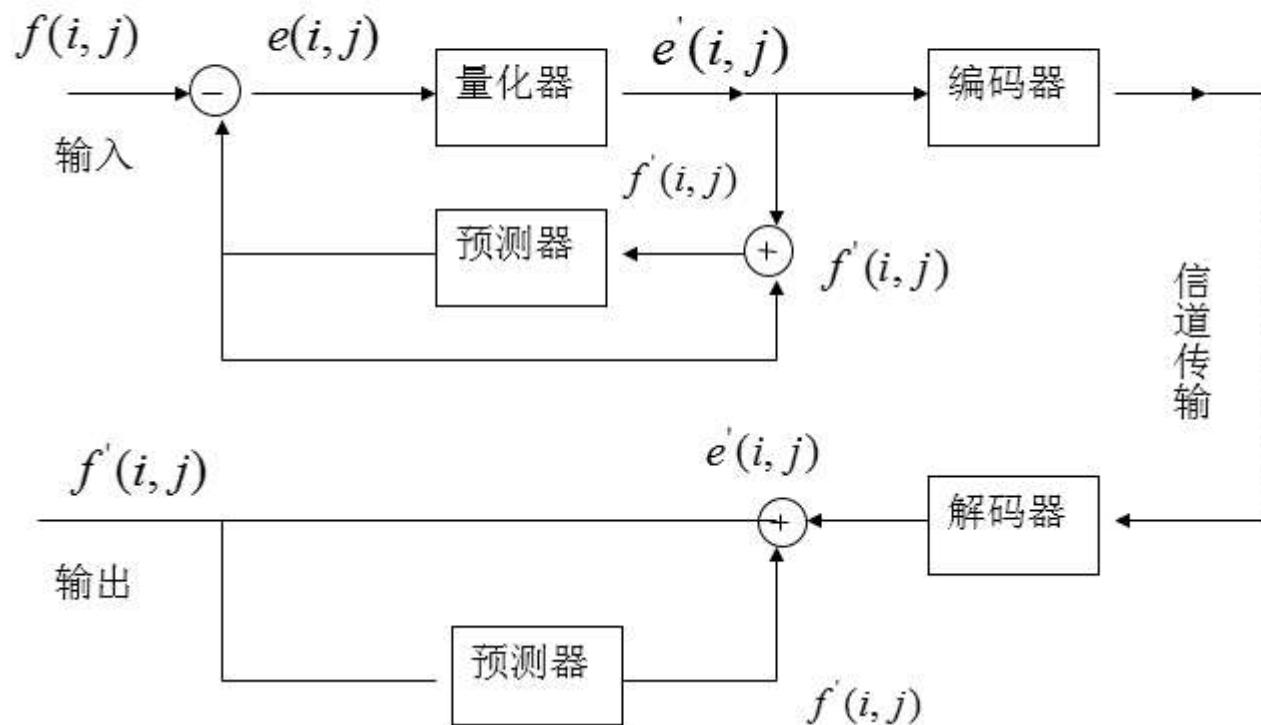


有损预测编码

- 在预测编码中，对差值信号进行量化后再进行编码就称之为有损预测编码。
- 有损预测方法有多种，常用的预测编码有：
△调制（Delta Modulation, 简称DM）和差分调制（Differential Pulse Code Modulation, DPCM）。DPCM是一种具有代表性的编码方法



DPCM系统由编码器和解码器组成，它们各有一个相同的预测器。**DPCM**系统的工作原理如下图所示：





- 系统包括发送、接收和信道传输三个部分。发送端由编码器、量化器、预测器和加减法器组成；接收端包括解码器和预测器等；信道传送以虚线表示。图中输入信号 $f(i, j)$ 是坐标 (i, j) 处的像素的实际灰度值，是由已出现先前相邻像素点的灰度值对该像素的预测灰度值。 $e(i, j)$ 是预测误差。
- DPCM包含量化器，这时编码器对编码，量化器导致了不可逆的信息损失，这时接收端经解码恢复出的灰度信号不是真正的 $f(i, j)$ ，而是重建信号。可见引入量化器会引起一定程度的信息损失，使图像质量受损。但是可以利用人眼的视觉特性，丢失不易觉察的图像信息，不会引起明显失真。



在图像压缩过程中，对数据的压缩通常真正体现在量化阶段。

量化器要完成的功能是按一定的规则对模拟取样值或数字图像像素值作近似表示，使经量化器输出的幅值个数有限，输出幅值的个数就称为量化器的级数。

由于量化器只能取有限多个量化级，量化过程是多对一的映射，因此不可避免地会引入误差，该误差称为量化误差，由于量化误差对图像的影响与噪声相似，有时也称量化误差为量化噪声。量化误差是造成压缩失真最主要的原因。量化过程是不可逆的，量化误差是不可恢复的，因此，加入量化器的图像压缩系统通常是有损压缩系统。



量化可以分为两大类：标量量化（又称为无记忆量化或一维量化）和向量量化（又称为记忆量化或多维量化）。

标量量化是针对每个取样值进行量化的，每个样值的量化都与其他样值无关，因此也称为无记忆量化。

向量量化不是针对单个取样值进行，而是每次对一组取样值进行量化，从码字集合中选出与输入样值序列失真最小的一个码字来近似该输入信号。

向量量化比标量量化具有更大的压缩能力。向量量化常与其他的编码方法结合使用，最常见的是与变换编码相结合使用。



DPCM 预测是一种近似处理系统，因为不同图像的自相关系数是不尽相同的，这一幅图像适用的模型和系数，对另一幅图像不一定适用。在实际工作中，针对不同情况，在预测模型上常采用相对固定的 a_i 系数的方法。



本章重点：

1. 图像压缩的概念，图像评价的方法
2. 图像的霍夫曼编码、算术编码、行程编码、预测编码、变换编码等方法
3. 图像压缩的标准

作业：第五章1, 2, 3, 4



图像压缩编码

山东科技大学
郑永果教授



第五章 图像压缩编

五

已讲解内容

1. 图像压缩的概念

图像压缩，“数据”与“信息”

数据冗余，编码冗余、像素间冗余、心理视觉冗余

无误差（亦称无失真、无损、信息保持）编码和有误差
(有失真或有损) 编码

图像编码模型，信源编码，信源解码，信道编码，信道解码

图像保真度准则，客观保真度准则，主观保真度准则

2. 统计编码方法：图像的霍夫曼编码、算术编码、

3. 无失真编码：位平面编码，二值图像编码，行程编码，无损预测编码

4. 有误差压缩编码：有损预测编码

本次课内容：

1. 正交变换编码：变换编码原理，卡胡南-列夫变换，离 散余弦变换

2. 图像编码的国际标准：静止图像压缩标准，运动图像压缩标准



5.5 正交变换编码

从理论上，采用正交变换不能直接对图像数据进行有效的压缩，但正交变换改变了图像数据的表现形式，为编码压缩提供了可能。

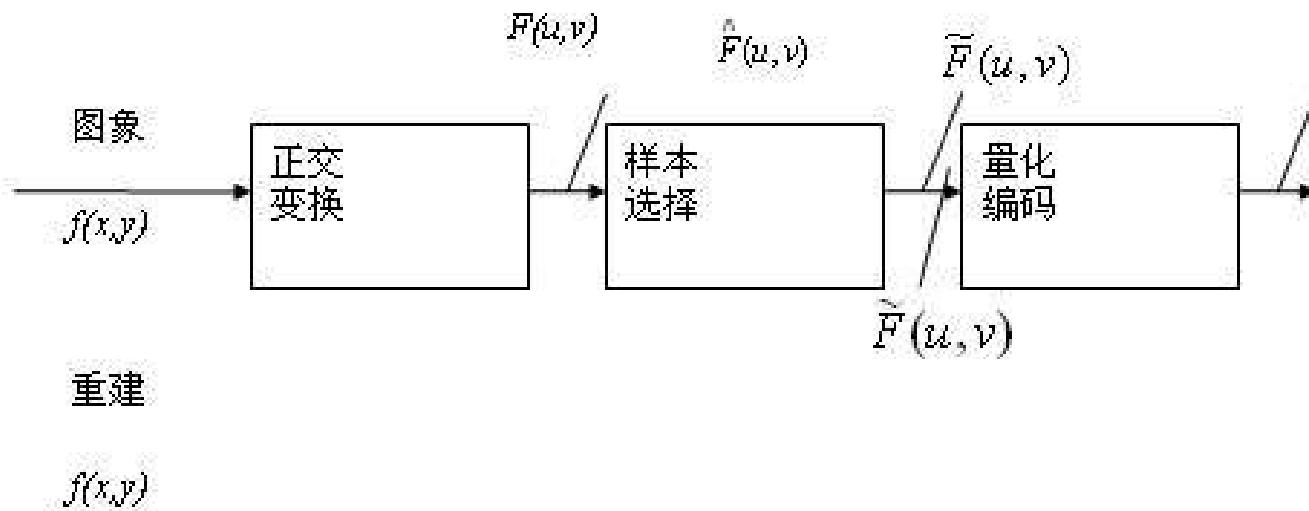


5.5.1 变换编码原理

变换编码不是直接对空域图像信号编码，而是首先将图像数据经过某种正交变换（如傅立叶变换（DFT），离散余弦变换（DCT），K-L变换等等）另一个正交矢量空间(称之为变换域)，产生一批变换系数，然后对这些变换系数进行编码处理，从而达到压缩图像数据的目的。



➤ 变换编码的原理 如下图：



➤ 图像数据经过正交变换后，空域中的总能量在变换域中得到保持，但像素之间的相关性下降，能量将会重新分布，并集中在变换域中少数的变换系数上，因此，选择少数 $F(u,v)$ 来重建图像就可以达到压缩数据的目的，并且重建图像仅引入较小误差。



思路：一幅 $N \times N$ 图像先被分割为 $n \times n$ 的子图像，通过变换这些子图像得到 $(N/n)^2$ 个 $n \times n$ 个子图像变换数组。变换的目的是解除每个子图像内部像素之间的相关性或将尽可能多的信息集中到尽可能少的变换系数上。量化时有选择性地消除或粗糙地量化携带信息最少的系数，因为它们对重建的子图像的质量影响最小。最后，对量化了的系数进行编码（常利用变长编码）。



解码部分由与编码部分相反排列的一系列逆操作模块构成。由于量化是不可逆的，所以解码部分没有对应的模块。



5.5.2 正交变换的性质

正交变换具有如下的性质：

- (1) 正交变换是熵保持的，说明正交变换前后不丢失信息。因此用图像各像素灰度存储或传送和用变换系数去存储或传输一样。
- (2) 正交变换熵能量保持的。
- (3) 正交变换重新分配能量。常用的正交变换，如傅立叶变换，能力集中于低频区，在低频区变换系数能量大而高频区系数能力小得多。这样可用



熵编码中不等长码来分配码长，能量大的系数分配较少的比特数，从而达到压缩的目的。同理，也可以用零代替能量较小的系数的方法压缩。

(4) 去相关性质。正交变换把空间域中高度相关的像素灰度值变为相关很弱或不相关得频率域系数。显然这样能去掉存在于相关性中的冗余度。



5.5.3 变换编码的数学分析

设A为正交矩阵，则有

$$Y = AX \quad (5-10)$$

由于A为正交矩阵，有

$$AA^T = AA^{-1} = E$$

传输或存储利用变换得到的Y，在接收端，经逆变换可恢复X

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

若在允许失真的情况下，传输和存储只用Y的前M(M<N)个分量，这样就得到Y的近似值：



$$\hat{Y} = [y_0, y_1, y_2, \dots, y_{M-1}]^T$$

来重建X，得到利用Y的近似值X的近似值为：

$$\hat{X} = A_1^T \hat{Y}$$

式中， A_1 为 $M \times M$ 阵。只要 A_1 选择恰当就可以保证重建图像的失真在一定允许限度内。关键的问题是如何选择A和 A_1 ，使之既能得到最大压缩又不造成严重失真。为此要研究X的统计性质。



对于 $X = [x_0, x_1, x_2, \dots, x_{n-1}]^T$

X 的均值为: $\bar{X} = E[X]$

X 的协方差矩阵为: $\Sigma_X = E[(X - \bar{X})(X - \bar{X})^T]$

同理, 对于 $Y = [y_0, y_1, y_2, \dots, y_{n-1}]^T$

Y 的均值为: $\bar{Y} = E[Y]$

Y 的协方差矩阵为: $\Sigma_Y = E[(Y - \bar{Y})(Y - \bar{Y})^T]$



根据式(5-10)得：

$$\begin{aligned}\Sigma_Y &= E[(AX - A\bar{X})(AX - A\bar{X})^T] \\ &= AE[(X - \bar{X})(X - \bar{X})^T]A^T \\ &= A\Sigma_X A^T\end{aligned}$$

可见， \mathbf{Y} 的协方差 Σ_Y 可由 Σ_X 作二维正交变换得到。 Σ_X 是图像固有的，因此关键是要选择合适的 A ，使变换系数 \mathbf{Y} 之间有更小的相关性。另外去掉了一些系数使得 $\hat{\mathbf{Y}}$ 误差不大。

总之，选择合适的 A 和相应的 A_1 ，使变换系数 \mathbf{Y} 之间的相关性全部解除和使 \mathbf{Y} 的方差高度集中，就称为最佳变换。



最佳变换A选择的准则：

若选择变换矩阵A，使 $\Sigma_{\hat{Y}}$ 为对角阵，那么变换系数之间的相关性可完全解除。接着选择集中主要能量的Y系数前M项，则得到的 \hat{Y} 将引起小的误差，使Y的截尾误差小。

最佳变换的核心在于经变换后能使 $\Sigma_{\hat{Y}}$ 为对角阵。

若采用某种变换矩阵A，变换后的 $\Sigma_{\hat{Y}}$ 接近于对角阵，则这种变换称为**准最佳变换**。



K-L变换是能满足均方误差准则下最佳变换。

K-L变换与图像数据有关，运算复杂，没有快速算法，因此K-L变换在使用性受到了很大的限制。

傅立叶变换、沃尔什变换、离散余弦变换，是常用的准最佳变换。



5.5.4 卡胡南 - 列夫变换 (K-L)

对于 $N \times N$ 的矩阵 T , 有 N 个标量 λ_i ,
 $i=1, 2, \dots, N$, 能使 $|T - \lambda_i I| = 0$ 则 λ_i 叫做矩阵 T 的特征值。

另外, N 个满足 $TV_i = \lambda_i V_i$ 的向量 V_i 叫做 T 的特征向量, 这些特征向量构成一个正交基集。

设 X 是一个 $N \times 1$ 的随机向量, 也就是说, X 的每个分量 x_i 都是随机变量。 X 的均值 (平均向量) 可以由 L 个样本向量来估计向量 M_x :

$$M_x \approx \frac{1}{L} \sum_{l=1}^L X_l \quad (5-11)$$



► M_x 协方差矩阵可以由

$$\Phi_{M_x} = E\{(X - M_x)(X - M_x)^T\} \approx \frac{1}{L} \sum_{l=1}^L X_l X_l^T - M_l M_l^T \quad (5-12)$$

来估计。协方差矩阵是实对称的。对角元素是个随机变量的方差，非对角元素是它们的协方差。

- 定义一个线性变换 T ，它可由任何 X 向量产生一个新向量 Y ：

$$Y = T(X - M_x) \quad (5-13)$$

- 式中， T 的各行是 M_x 的特征向量，即 T 的行向量就是 M_x 的特征向量。



- 变换得到的Y是期望为零的随机向量。Y的协方差矩阵可以由X的协方差矩阵决定：

$$\Phi_Y = T \Phi_X T^T \quad (5-14)$$

因为T的各行是 Φ_X 的特征向量，故 Φ_Y 是一个对角阵，对角元素是 Φ_X 特征值。因此

$$\Phi_Y = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \lambda_N \end{bmatrix}$$

这些也是的 Φ_X 特征值。

- 随机向量Y是由互不相关的随机变量组成的，因此线性变换T起到了消除变量间的相关性的作用。



图像压缩编码

山东科技大学
郑永果教授



- 特征向量变换是可逆的。
- 要实现对信号进行K—L变换，首先要求出矢量x的协方差矩阵 Φ_x ，再求协方差矩阵 Φ_x 的特征值 λ_i ，然后求 λ_i 对应的 Φ_x 的特征向量，再用 Φ_x 的特征向量构成正交矩阵T。
- 例：若已知随机矢量x的协方差矩阵为

$$\Phi_x = \begin{bmatrix} 6 & 2 & 0 \\ 2 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

求其正交矩阵T？



1) 按 $|A - \Phi x| = 0$, 求 Φx 的特征值 λ_i :

$$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} - \begin{bmatrix} 6 & 2 & 0 \\ 2 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} = 0 \text{ 得: } \begin{bmatrix} \lambda - 6 & -2 & 0 \\ -2 & \lambda - 2 & 1 \\ 0 & 1 & \lambda - 1 \end{bmatrix} = 0$$

则可解得: $\lambda_1 = 6.854$ $\lambda_2 = 2$ $\lambda_3 = 0.146$

2) 求 λ_i 对应的特征向量。利用 $\lambda 1$, $\lambda 2$, $\lambda 3$ 分别求得如下三个特征向量:

$$V_1 = \begin{bmatrix} 0.918 \\ 0.392 \\ -0.067 \end{bmatrix} \quad V_2 = \begin{bmatrix} 0.333 \\ -0.667 \\ 0.667 \end{bmatrix} \quad V_3 = \begin{bmatrix} -0.217 \\ 0.634 \\ 0.742 \end{bmatrix}$$



用V1, V2, V3的转置向量作为正交矩阵T的行向量，
那么，对于任一均值为0的向量X=(2, 1, -0.1)的KL变
换为：

$$Y = TX = \begin{bmatrix} 0.918 & 0.329 & -0.067 \\ 0.333 & -0.667 & 0.667 \\ -0.217 & 0.634 & 0.742 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -0.1 \end{bmatrix} = \begin{bmatrix} 2.234 \\ -0.067 \\ 0.127 \end{bmatrix}$$

则Y的协方差矩阵Φ_Y为：

$$\Phi_Y = T\Phi_X T^T = \begin{bmatrix} 6.854 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0.146 \end{bmatrix}$$



5.5.5 离散余弦变换 (DCT)

- 在数字图像压缩编码中，最佳变换 K-L 计算复杂，一般不采用。
- 由于 DCT 与 K-L 变换压缩性能和误差很小，接分年近而 DCT 计算复杂度适中，又具有以近离特性，在图像数据压缩中，采用离散余弦变换编码的方案很多。
- JPEG、MPEG、H.261 等压缩标准，都用到离散余弦变换编码进行数据压缩。
- 余弦变换是傅立叶变换的一种特殊情况。函在数包是实偶函数，再将离散余弦变换级数含余弦项，称之为离散余弦变换 DCT (Discrete Cosine Transform)。





➤ 二维离散偶余弦正变换公式为：

$$C(u,v) = E(u)E(v) \frac{2}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cdot \cos\left(\frac{2x+1}{2N} u\pi\right) \cdot \cos\left(\frac{2y+1}{2N} v\pi\right) \quad (5-15)$$

式中， $x, y, u, v = 0, 1, \dots, N-1$ 。

$$E(u)E(v) = 1/\sqrt{2} \quad \text{当 } u=v=0 \text{ 时。}$$

$$E(u)E(v) = 1 \quad \begin{aligned} &\text{当 } u=1, 2, \dots, N-1; \\ &\text{v}=1, 2, \dots, N-1 \text{ 时。} \end{aligned}$$



➤ 二维离散偶余弦逆变换公式为：

$$f(x, y) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} E(u)E(v)C(u, v) \cdot \cos\left(\frac{2x+1}{2N}u\pi\right) \cdot \cos\left(\frac{2y+1}{2N}v\pi\right) \quad (4-39)$$

式中 $x, y, u, v = 0, 1, \dots, N-1$ 。

$$E(u)E(v) = 1/\sqrt{2} \quad \text{当 } u=v=0 \text{ 时。}$$

$$E(u)E(v) = 1 \quad \text{当 } u=1, 2, \dots, N-1; \\ v=1, 2, \dots, N-1 \text{ 时。}$$



- 二维离散余弦变换核具有可分离特性，所以，其正变换和逆变换均可将二维变换分解成系列一维变换(行、列)进行计算。
- 在DCT为主要方法的变换编码中，一般不直接对整个图像进行变换，而是首先对图像分块，将 $M \times N$ 的一幅图像分成不重叠的 $M/K \times N/K$ 个 $K \times K$ 块分别进行变换。
 - 这样做好处主要体现在：
 - 第一，降低运算量，如对一幅 512×512 图像，分块变换仅需约 $1/3$ 的运算量；
 - 其次，后续的量化和扫描处理可以得到明显的简化；
 - 第三，容易将传输误差引起的错误控制在一个块内，而不是整个图像扩散。分块大小通常选 8×8 和 16×16 。



5.5.6 编码

变换为压缩数据提供了条件，压缩数据还要通过编码来实现。

通常所用的编码方法有两种：一是区域编码法；二是门限法。



1. 区域编码法

这种方法的关键在于选出能量集中的区域。

例如，正交变换后变换域中的能量多半集中在低频率空间上，在编码过程中就可以选取这一区域的系数进行编码传送，而其他区域的系数可以舍弃不用。在解码端，可对舍弃的系数进行补零处理。

这样，由于保持了大部分图像能量，在恢复图像中带来的质量劣化并不显著。



区域编码法的缺点：

一旦选定某个区域就固定不变了，有时图像中的能量也会在其它区域集中较大的数值，舍弃它们会造成图像质量较大的损失。



2. 门限编码法

门限编码法，不是选择固定的区域，而是事先设定一个门限值 T ，如果系数超过 T 值，就保留下来并且进行编码传送。如果系数值小于 T 值就舍弃不用。

这种方法有一定的自适应能力。它可以得到较区域编码好的图像质量。



但是这种方法也有缺点，那就是超过门限值的系数的位置是随机的。因此，在编码中除对系数值编码外，还要有位置码。这两种码同时传送才能在接收端正确恢复图像。所以，其压缩比有时会有所下降。



5.6 图像编码的国际标准简介

国际标准主要是由国际标准化组织（ISO）和国际电信联盟（ITU）制定的。国际电信联盟的前身是国际电话电报咨询委员会（CCITT）。

国际标准主要包括三个部分：

静止灰度（彩色）图像压缩标准（JPEG）。

运动图像压缩标准（MPEG）。

和二值图像压缩标准（JBIG）。



- JPEG是联合图像专家小组开发研制的连续色调、多级灰度、静止图像的数字图像压缩编码方法。JPEG中的核心算法是DCT变换编码。
- MPEG视频压缩分为空间域压缩与时间域压缩。
- JBIG(Joint Bi-level Image Experts Group, 联合二值图像专家组)为二值图像编码标准。



5.6.1 静止图像压缩标准

由上述两个组织组成的联合静止图像专家组 JPEG (Joint Photographic Experts Group, 简称 JPEG) 建立了静止灰度 (或彩色) 图像压缩的公开算法, 并于1991年开始使用。这种压缩算法成为国际上通用的标准, 因此又称为JPEG标准。

JPEG标准适于静图像的压缩, 电视图像序列的帧内图像的压缩编码也常采用JPEG压缩标准。



JPEG专家组定义了三种编码系统：

- (1) DCT有损编码系统；
- (2) 扩展编码系统；
- (3) 无失真编码系统。

在视觉效果不受到严重损失的前提下，JPEG算法可以达到15到20的压缩比。如果在图像质量上稍微牺牲一点的话，可以达到40：1或更高的压缩比。



如果处理的是彩色图像，JPEG算法首先将RGB分量转化成亮度分量和色差分量，同时丢失一半的色彩信息（空间分辨率减半）；然后，用离散余弦变换来进行变换编码，舍弃高频的系数，并对余下的系数进行量化，以进一步减少数据量。最后，使用行程长度编码和Huffman编码来完成压缩任务。



近年来，JPEG专家组又制定了**JPEG2000**标准，**JPEG2000**与传统**JPEG**最大的不同在于：它放弃了**JPEG**所采用的以离散余弦变换为主的区块编码方式，而改用以小波变换为主的多解析编码方式。



1 、 JPEG 的工作模式

JPEG对每一个图像分量单独编码。

JPEG对每个不同的图像分量可以采用不同的量化参数和熵编码的码表

对于一个图像分量， JPEG提供4种工作模式。



- **顺序编码：**

每一个图像分量按从左到右，从上到下扫描，一次扫描完成编码。

- **累进编码：**

图像编码在多次扫描中完成。

- **无失真编码：**

解码后能精确地恢复源图像采样值，其压缩比低于有失真压缩编码方法。

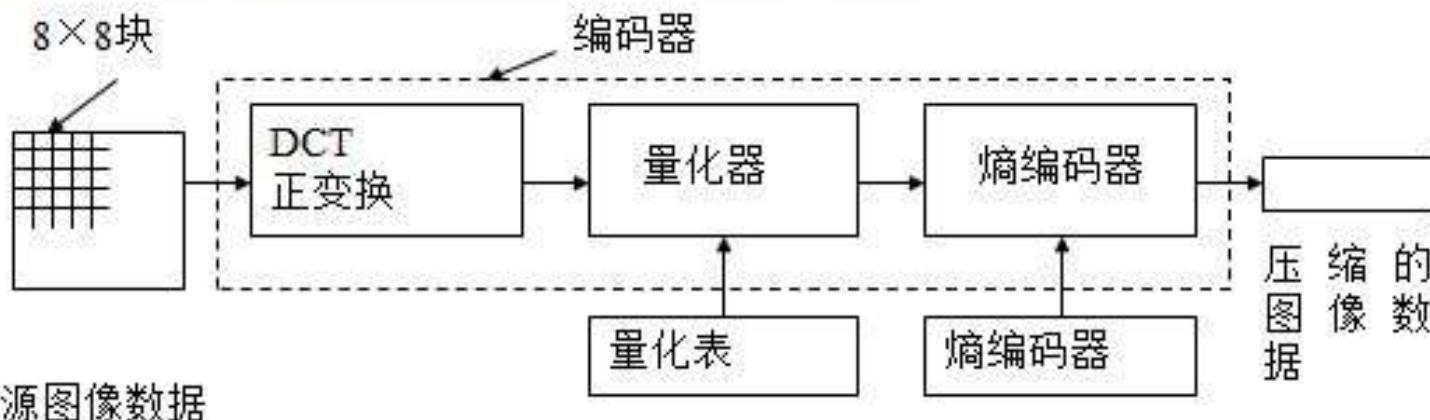
- **分层编码：**

图像在多个空间分辨率进行编码。

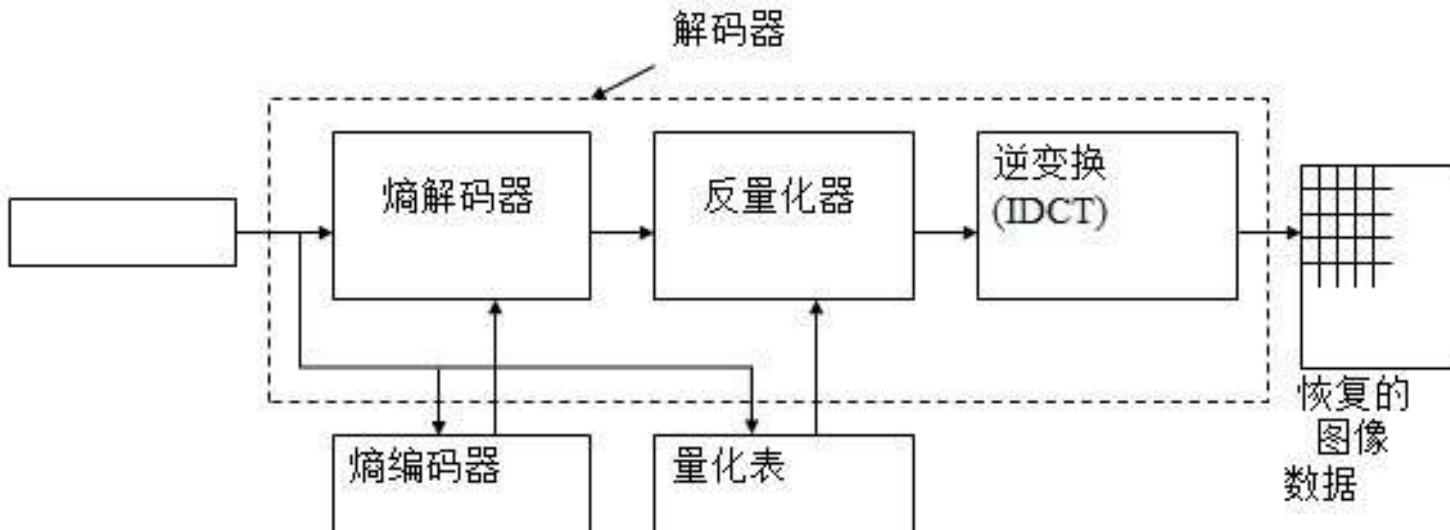


2 基本工作模式

基于DCT JPEG编码的过程框图



- JPEG采用的是 8×8 大小的子块的二维离散余弦变换(DCT)。
- 在编码器的输入端，把原始图像顺序地分割成一系列 8×8 的子块，设原始图像的采样精度为P位，是无符号整数，输入时把 $[0, 2^P]$ 范围的无符号整数变成 $[-2^{P-1}, 2^{P-1}-1]$ 范围的有符号整数，以此作为离散余弦正变换的输入。



- 在解码器的输出端经离散余弦逆变换(**IDCT**)后，得到一系列 8×8 的图像数据块，需将其数值范围由 $[-2^{P-1}, 2^{P-1}-1]$ 再变回到 $[0, 2^P]$ 范围内的无符号整数，来获得重构图像。



- 为了达到压缩数据的目的，对DCT系数需作量化处理。量化处理是一个多到一的映射，它是造成DCT编解码信息损失的根源。
- 在JPEG中采用线性均匀量化器，量化定义为对64个DCT系数除以量化步长，四舍五入取整。
- 量化的作用是在一定的主观保真度图像质量前提下，丢掉那些对视觉效果影响不大的信息。



例：给定Lena图像的一个平坦区域（ 8×8 子块）如下：

69	71	75	79	84	89	91
69	70	73	76	83	90	95
77	74	76	74	85	89	95
71	73	76	79	86	91	93
74	77	77	82	88	91	93
78	76	80	84	88	92	95
76	78	80	85	93	94	95
74	79	81	85	86	94	94

给出DCT变换单元量化过程。



- 用JPEG的亮度量化矩阵式对每个系数进行均匀量化，量化器输出为：



➤ 反量化后，进行 DCT 反变换，得到的解码图像为：

80	75	71	72	78	85	89	90
80	75	71	72	78	85	89	90
80	76	72	73	79	86	90	91
81	77	72	74	80	87	91	92
82	77	73	74	81	87	91	93
83	78	74	75	81	88	92	93
83	79	75	76	82	89	93	94
84	79	75	76	82	89	93	94



➤如下是它的DCT变换系数，可以看到能量集中在少数低频系数：

660.1250	-47.0496	25.9980	10.3993	7.8750	8.4866	5.6025	1.3176
-17.3267	-2.6749	5.2236	-1.3234	0.5222	0.2914	0.2800	-2.281
0.0280	-0.6463	-0.9545	0.9620	2.4730	1.9783	-0.316	2.1741
2.3003	0.4542	-2.2403	3.5559	1.2907	-1.0024	0.1580	0.9747
-2.3750	0.1038	-3.2220	0.9653	1.3750	2.2258	0.3875	3.5236
0.9294	-1.3282	-2.4256	0.9828	-1.9317	-0.6972	0.1253	-1.856
0.3943	2.6640	-0.5669	-3.4168	-0.8891	-1.6182	-2.545	-1.732
2.1666	1.7238	-0.3335	-0.4808	-2.6253	-0.9699	1.4854	-1.183



5.6.2 运动图像压缩标准

- 从时间的观点看，数字图像分为静态图像和运动图像，视频信号就是典型的运动图像。
- 视频压缩的目标是在尽可能保证视觉效果的前提下减少视频数据率。
- 根据压缩前和解压缩后的数据是否完全一致，视频压缩可分为有损压缩和无损压缩。
- 无损压缩意味着解压缩后的数据与压缩前的数据完全一致。有损压缩则意味着解压缩后的数据与压缩前的数据不一致。



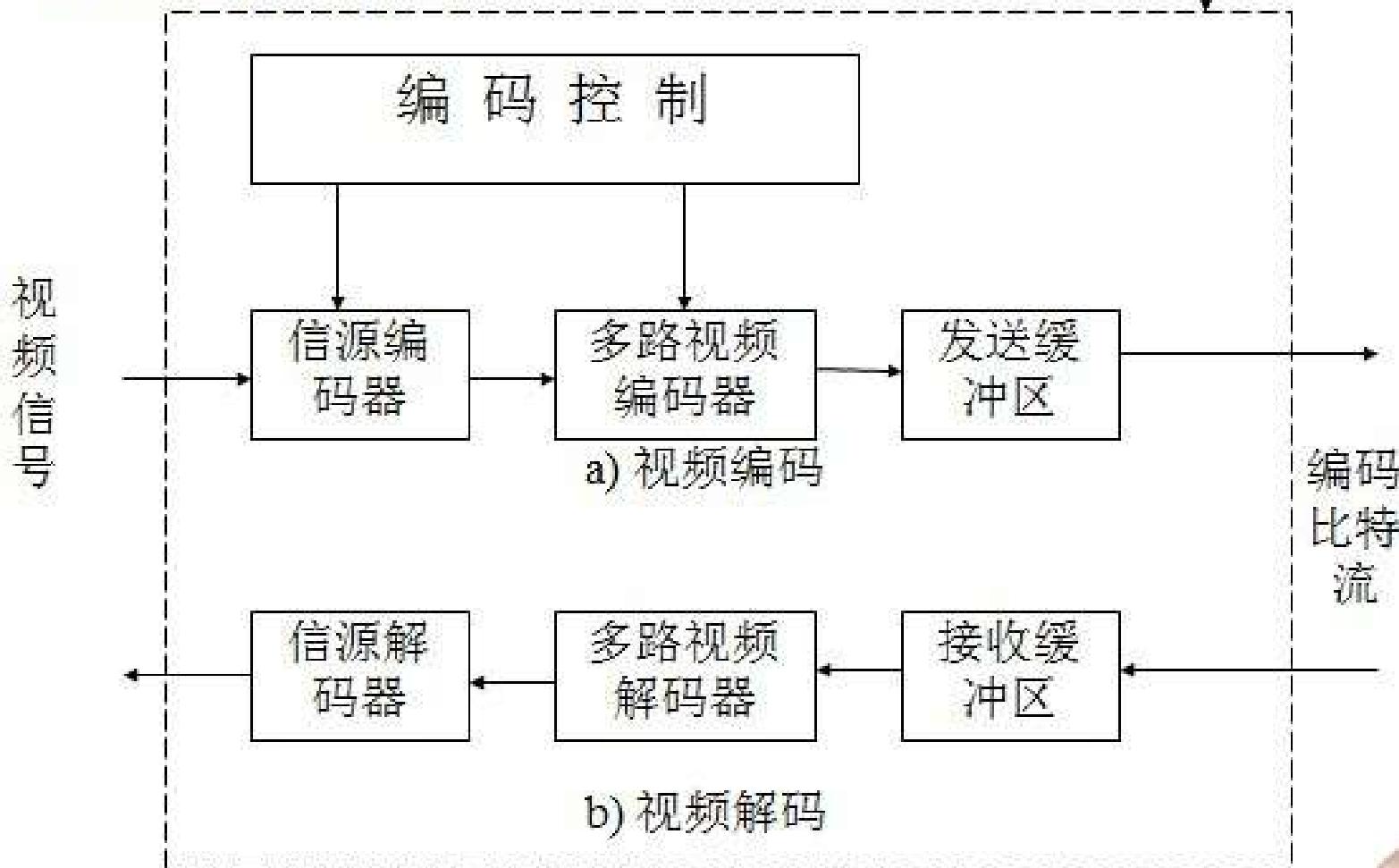
MPEG是运动图像专家组（Moving Photographic Experts Group）的简称，他们的任务是制定用于数字存储媒介中活动图像及伴音的编码标准。

MPEG与JPEG算法在概念上类似，只不过它还利用了相继图像之间的冗余信息。由于可以达到100：1的压缩比，所以MPEG算法非常实用。



外部控制

➤ 视频编解码过程：





- 视频信号的压缩包括两个主要方面：帧内压缩与帧间压缩。
- 帧内（**Intra frame**）压缩也称为空间压缩（**Spatial compression**）。当压缩一帧图像时，仅考虑本帧的数据而不考虑相邻帧之间的冗余信息。
- 帧间（**Inter frame**）压缩是基于许多视频或动画的连续前后两帧具有很大的相关性。即连续的视频其相邻帧之间具有冗余信息。根据这一特性，压缩相邻帧之间的冗余量就可以进一步提高压缩量，减小压缩比。帧间压缩也称为时间压缩（**Temporal compression**），它通过比较时间轴上不同帧之间的数据进行压缩。



- MPEG (Moving Picture Expert Group)是运动图像专家组的简称。该小组于1991年底提出了用于数字存储媒介的、速率约1.5MB / s的运动图像及其伴音的压缩编码，并于1992年正式通过，通常被称为MPEG标准，此标准后来被定名为MPEG-1。
- 到目前为止，MPEG标准已不再是一个单一的标准，而是一个用于全运动视频和相关音频压缩的标准系列，包括MPEG-1、MPEG-2、MPEG-3、MPEG-4和MPEG-7共5个标准，每一个标准都有其特定的应用范围。其中，MPEG-1和MPEG-2标准的应用范围最广。



- MPEG-1用于加速CD-ROM中图像的传输。
- MPEG-2用于宽带传输的图像，图像质量达到电视广播甚至HDTV的标准。和MPEG-1相比，MPEG-2支持更广的分辨率和比特率范围，将成为数字图像盘（DVD）和数字广播电视的压缩方式。
- MPEG-4标准支持非常低的比特率的数据流的应用，如电视电话，视频邮件和电子报刊等。



MPEG-7

MPEG-7(它的由来是1 2 4=7，因为没有MPEG-3、MPEG-5、MPEG-6)于1996年10月开始研究。确切来讲，MPEG-7并不是一种压缩编码方法，其正规的名字叫做“多媒体内容描述接口”，其目的是生成一种用来描述多媒体内容的标准，这个标准将对信息含义的解释提供一定的自由度，可以被传送给设备和电脑程序，或者被设备或电脑程序查取。

MPEG-7并不针对某个具体的应用，而是针对被MPEG-7标准化了的图象元素，这些元素将支持尽可能多的各种应用。建立MPEG-7标准的出发点是依靠众多的参数对图象与声音实现分类，并对它们的数据库实现查询。



- MPEG - MPEG—21

MPEG在1999年10月的MPEG会议上提出了“多媒体框架”的概念，同年的12月的MPEG会议确定了MPEG-21的正式名称是“多媒体框架”或“数字视听框架”，它以将标准集成起来支持协调的技术以管理多媒体商务为目标，目的就是理解如何将不同的技术和标准结合在一起需要什么新的标准以及完成不同标准的结合工作。



- MPEG视频压缩分为空间域压缩与时间域压缩。
- MPEG标准在空间域的压缩，每一帧被作为独立的图像获取，且压缩步骤与JPEG标准的步骤一样。时间域压缩，即帧间编码的基本思想是仅存储运动图像从一帧到下一帧的变化部分，而不是存储全部图像数据，这样做能极大地减少运动图像数据的存储量。
- 这是通过把帧序列划分成I帧、P帧、B帧，使用参照帧及运动补偿技术来实现的。



总结

- 本章在分析图像编码的必要性与可能性的基础上，对图像编码与压缩的基本概念、理论及其编码分类进行了简要介绍。并从无损压缩和有损压缩的角度具体介绍了几种常用的图像编码与压缩技术。最后介绍了图像编码标准。

1. 图像压缩的概念

图像压缩，“数据”与“信息”

数据冗余，编码冗余、像素间冗余、心理视觉冗余

无误差（亦称无失真、无损、信息保持）编码和有误差
(有失真或有损) 编码

图像编码模型，信源编码，信源解码，信道编码，信道
解码

图像保真度准则，客观保真度准则，主观保真度准则



- 无损压缩是指可以精确无误地从压缩数据中恢复出原始数据的图像压缩方法。
- 统计编码方法：图像的霍夫曼编码、算术编码、无失真编码：位平面编码，二值图像编码，行程编码，无损预测编码



- 有损编码是以丢失部分信息为代价来换取高压缩比的。有损压缩方法主要有损预测编码方法、变换编码方法等。

图像编码的国际标准

- JPEG是联合图像专家小组开发研制的连续色调、多级灰度、静止图像的数字图像压缩编码方法。JPEG中的核心算法是DCT变换编码。
- MPEG视频压缩分为空间域压缩与时间域压缩。



作业： P165-166 第五章 1, 2, 3, 4, 5



数字图像处理

(Digital Image Processing)

山东科技大学信息与电气工程学院

范 迪



第 6 章 图像分割

6.1 概述

6.2 阈值分割法

6.3 边缘检测分割法

6.4 区域分割法

6.5 Hough 变换

6.6 基于形态学的分割



6.1 概述

- 1、为什么要进行图像分割？
- 2、分割在图像分析系统中处于什么位置？
- 3、图像分割的概念
- 4、图像分割的基本思路
- 5、图像分割的基本方法



1、为什么要进行图像分割？

分割就是要把图像分成互不重叠的区域并提取感兴趣的目标

1) 汽车车牌识别



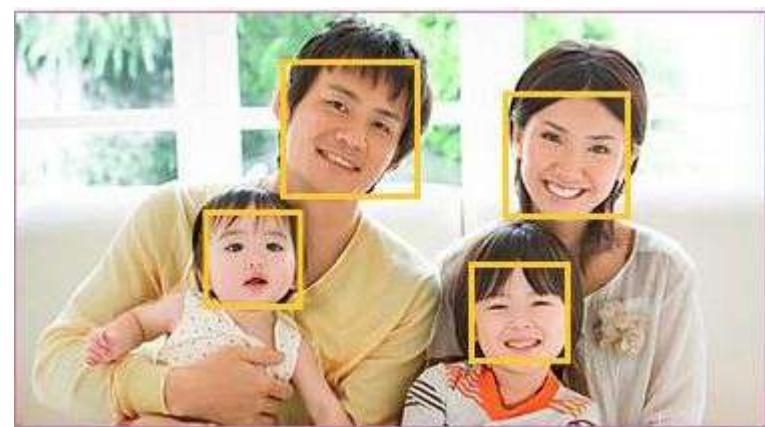
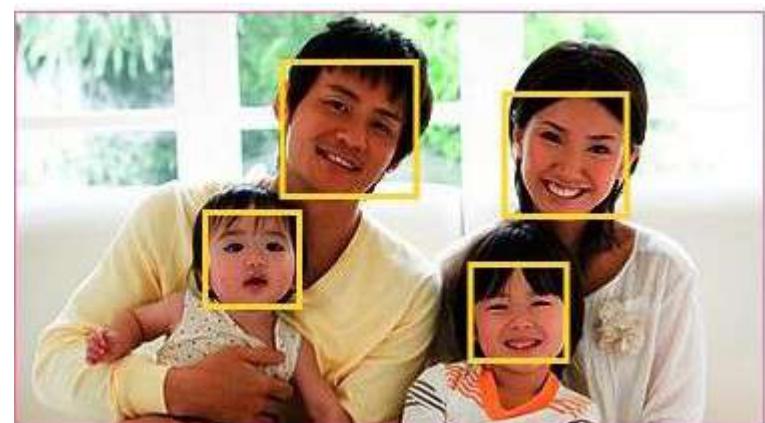
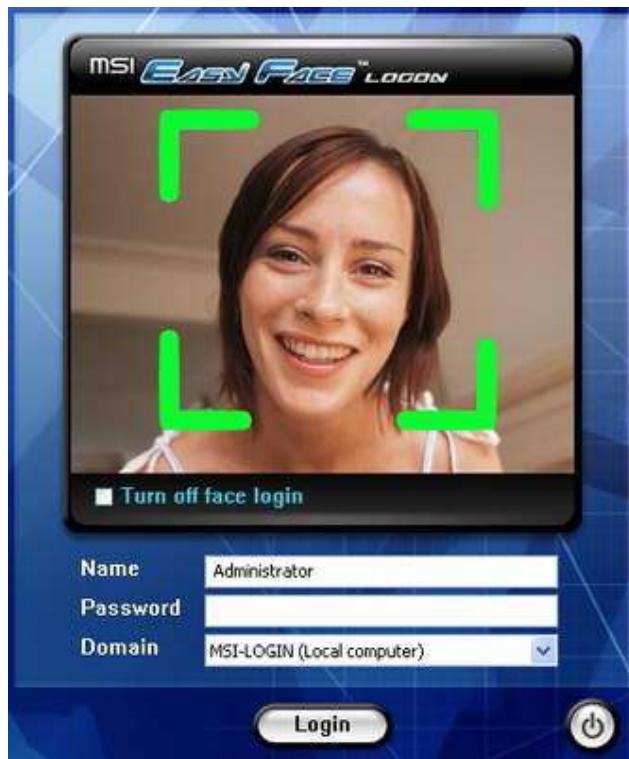
地点：托依堡 车号：N21055 时间：2008-07-15 16:27:16 速度：33km/h 限速：80km/h 类型：正常





1、为什么要进行图像分割？

2) 人脸检测和识别





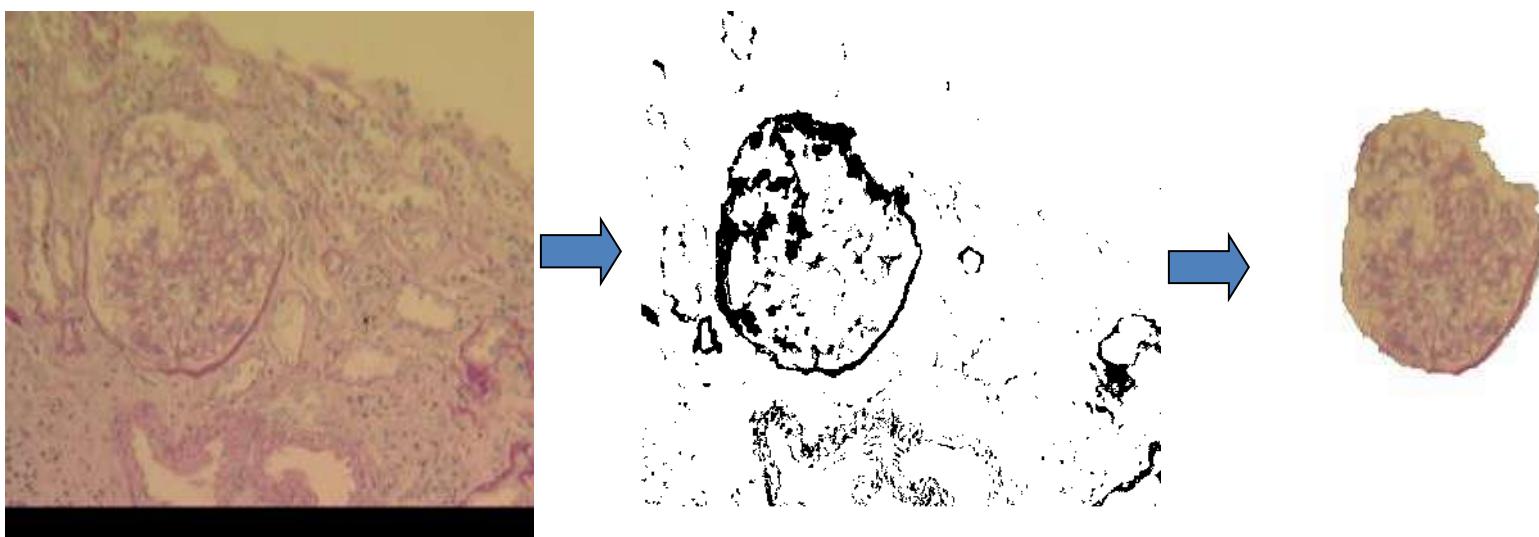
1、为什么要进行图像分割？

3) 产品包装字符文字识别





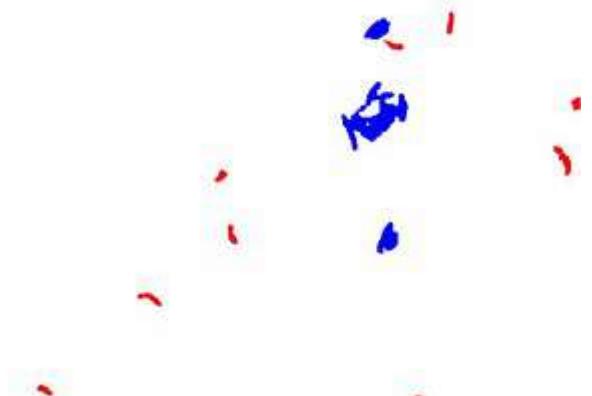
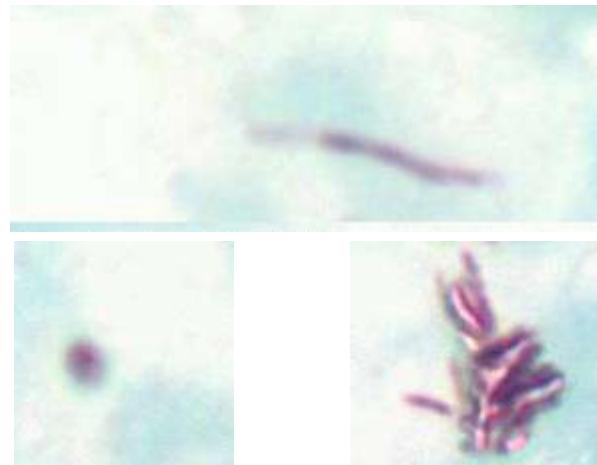
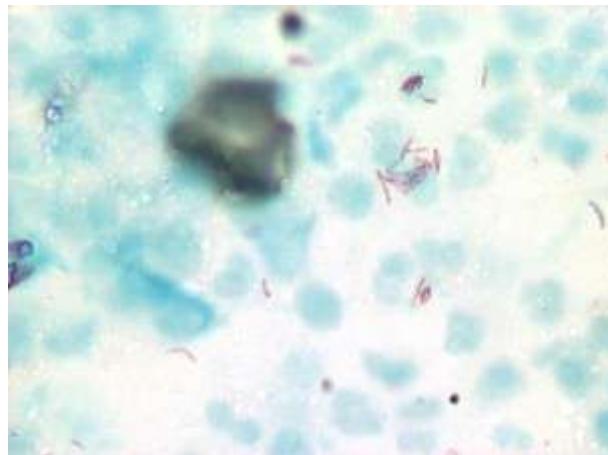
4) 肾组织切片图像的处理与识别



通过图像分割技术，获得肾小球区域的边界（闭合），最终对所提取出的肾小球内部的细胞核进行定性与定量分析。



5) 细菌自动检测分析

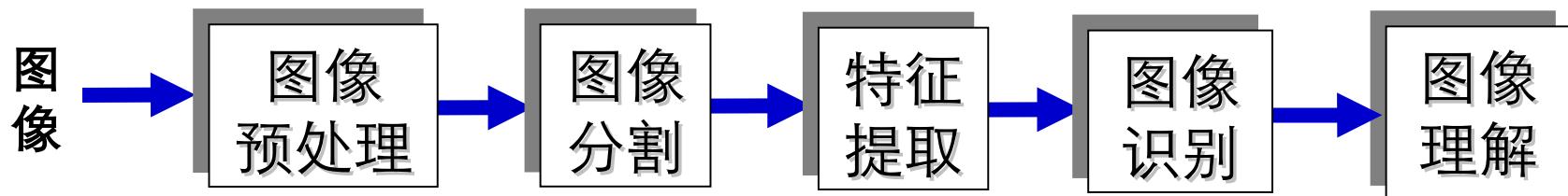


结果：单细菌 9 个，聚团细菌 3 个

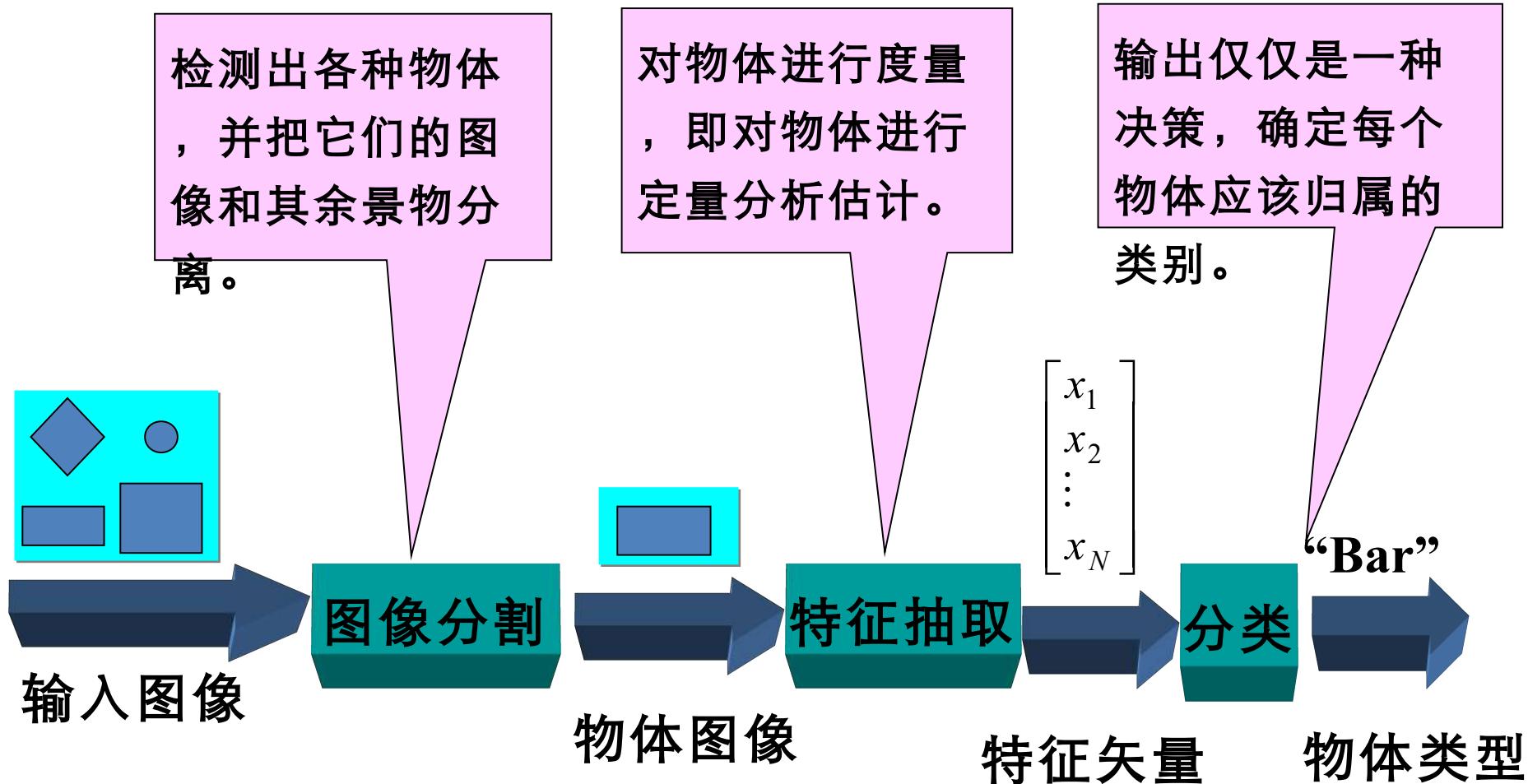
自动对显微镜下的细菌进行检测与统计，可以抵抗染色以及杂质的影响。



图像的处理过程为：

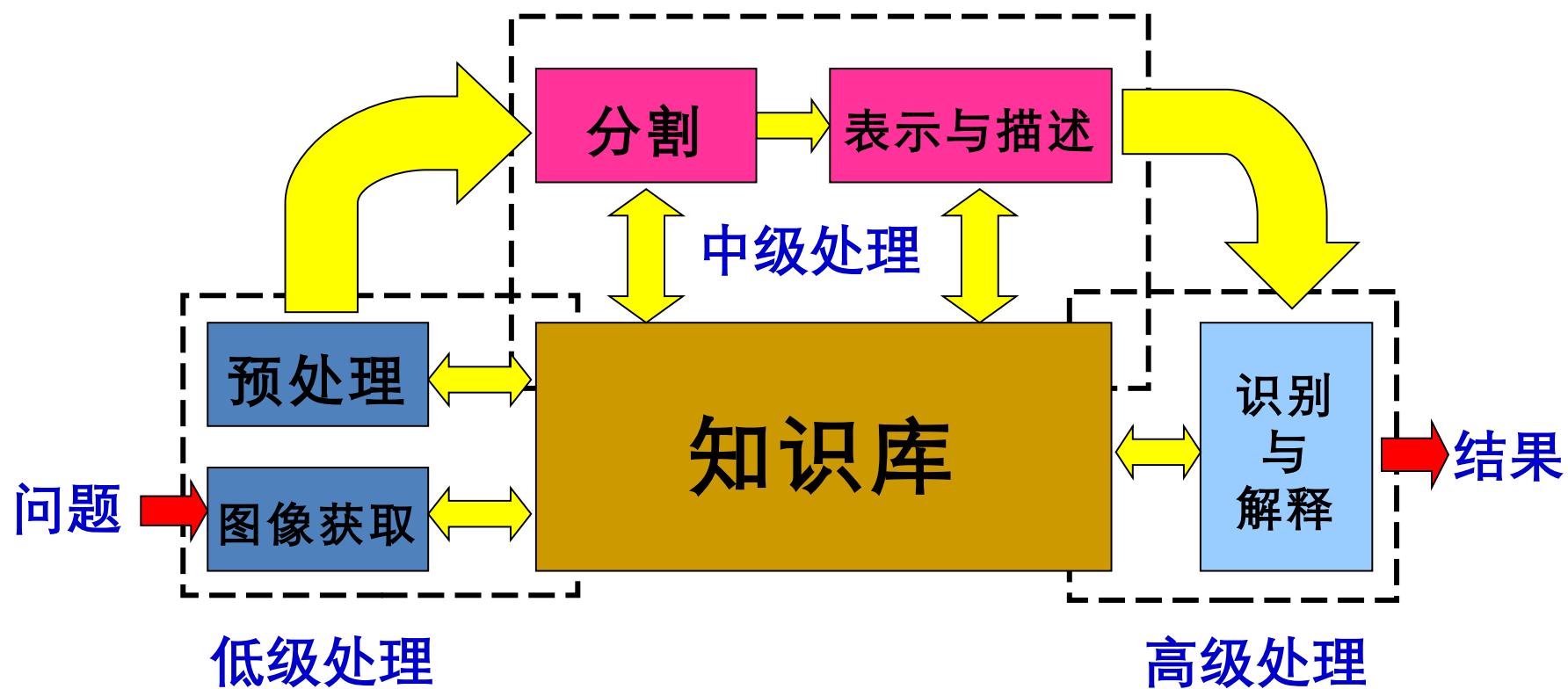


有多个物体的数字图像，其处理过程分为三个阶段：





2、分割在图像分析系统中处于什么位置？



图像分析系统的基本构成



**图像分割是图像处理过渡到图像分析的关键步骤
，也是一种基本的计算机视觉技术。**



3、图像分割的概念

定义：把图像分成各具特性的互不重叠的区域，并提取出感兴趣目标的技术和过程。

可以是灰度、颜色、纹理等

对应单个区域和多个区域

还可理解为：将整幅图像区域 R 划分成 n 个子区域，满足的条件为：

$$1) \bigcup_{i=1}^n R_i = R$$

$$2) R_i \cap R_j = \emptyset, \text{ 对所有的 } i \text{ 和 } j, i \neq j$$

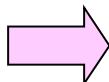
$$3) P(R_i) = \text{TRUE}, i = 1, 2, \dots, n$$

$$4) P(R_i \cup R_j) = \text{FALSE}, i \neq j$$

$$5) R_i \text{ 是一个连通的区域}, i = 1, 2, \dots, n$$

定义包含的意义

1) $\bigcup_{i=1}^n R_i = R$



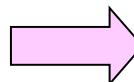
全部子区域的总和（并集）应能包括图像中所有像素，或者说分割应将图像中的每个像素都分进某 1 个子区域中。

2) 对所有的 i 和 j , $i \neq j$, 有 $R_i \cap R_j = \emptyset$



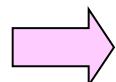
各个子区域互不重叠，或者说 1 个像素不能同时属于 2 各区域。

3) $i=1,2\dots n$,
有 $P(R_i)=\text{TRUE}$



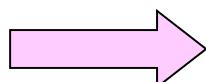
属于同一区域的像素应该具有某些相同特性，区域边界是明确的。

4) 对 $i \neq j$, 有
 $P(R_i \cup R_j)=\text{FALSE}$



属于不同区域中的像素应该具有一些不同的特性。

5) 对 $i=1,2\dots n$,
 R_i 是连通的区域



要求同 1 个子区域内的像素应当是连通的。

分割准则应适用于所有区域和像素

分割准则应能确定各区域像素有代表性的特性。



4、图像分割的基本思路

- 1) 从简到难，逐级分割；
- 2) 控制背景环境，降低分割难度；
- 3) 把焦点放在增强感兴趣对象、减小不相干部分的干



感兴趣的对像：汽车牌照

形区域



不相干图像成分：非矩



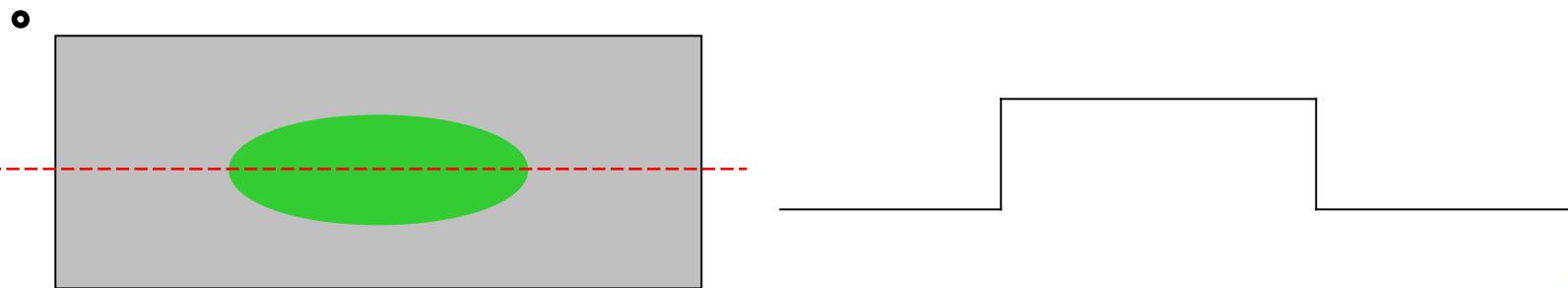


分割算法基于灰度值的两个基本特性：不连续性和相似性。

不连续性——区域之间

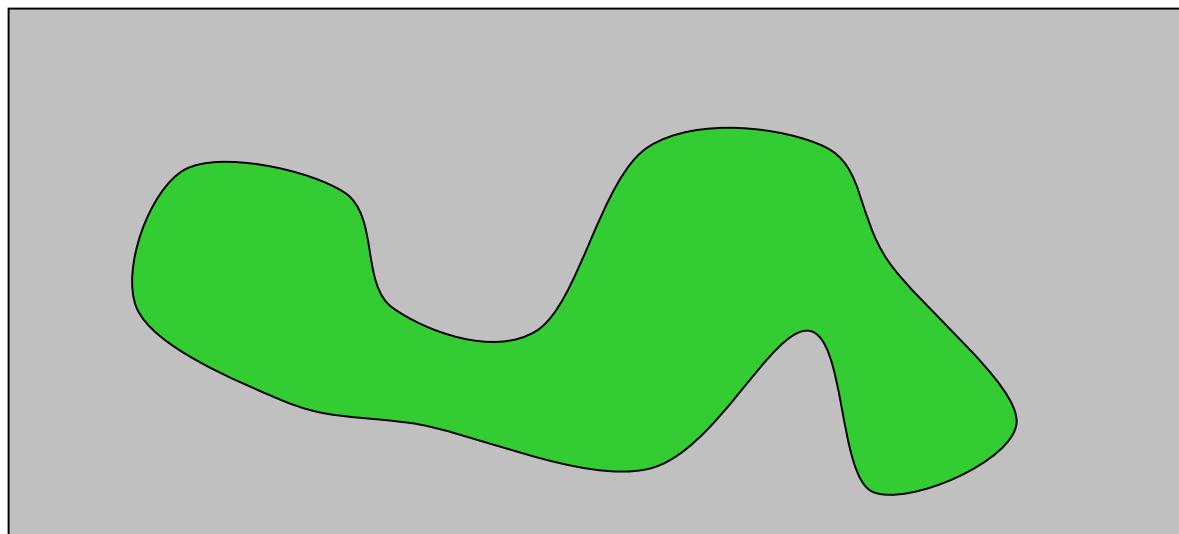
相似性——区域内部

作法 1：根据图像像素灰度值的**不连续性**，先找到点、线（宽度为 1）、边（不定宽度），再确定出区域





作法 2：根据图像像素灰度值的**相似性**，通过选择阈值，找到灰度值相似的区域，区域的外轮廓就是目标对象的边。





5、图像分割的基本方法

基于区域之间灰度不连续性的一类分割方法：基于边缘的分割方法，先提取区域边界，再确定边界限定的区域。

如：边缘检测分割法、Hough 变换等。

基于区域内部灰度相似性的一类分割方法：区域分割方法，确定每个像素的归属区域，从而形成一个区域图。

如：阈值分割法，区域生长、分裂合并、形态学分割等

。



对分割的要求：

- (1) **有效性**: 对各种分割问题有效的准则，能将感兴趣的区域或目标分割出来。
- (2) **整体性**: 即能得到感兴趣区域的封闭边界，该边界无断点和离散点。
- (3) **精确性**: 得到的边界与实际期望的区域边界很贴近。
- (4) ... 结果受噪声影响很小。



数字图像处理

(Digital Image Processing)

山东科技大学信息与电气工程学院

范 迪



6.2 阈值分割法

6.2.1 阈值分割原理

6.2.2 阈值选取方法



6.2.1 阈值分割原理

1、阈值分割的原理

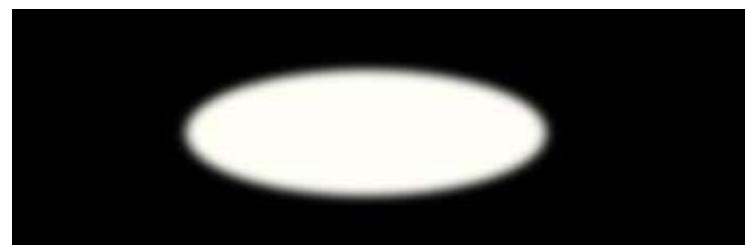
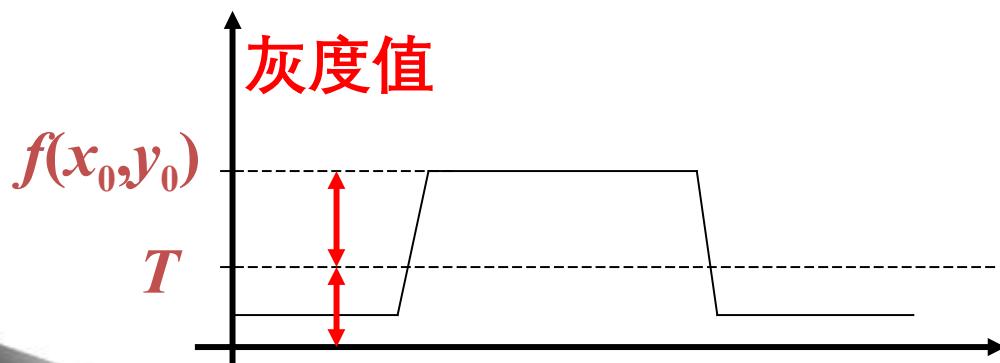
图像阈值分割是一种广泛应用的分割技术。

利用图像中要提取的**目标物**与其**背景**在**灰度特性**上的差异，把图像视为具有不同灰度级的两类区域（**目标**和**背景**）的组合，选取一个**合适的阈值**，以确定图像中每个**像素点**应该属于**目标**还是**背景区域**，从而产生相应的**二值图像**。



设原始图像 $f(x,y)$ ，以一定的准则在 $f(x,y)$ 中找出一个合适的灰度值，作为阈值 t ，分割后的图像 $g(x,y)$ 有以下几种形式(1) 设阈值为 t ，阈值化为：

$$g(x,y)=\begin{cases} 1 & f(x,y)\geq t \\ 0 & f(x,y)<t \end{cases} \quad \text{或} \quad g(x,y)=\begin{cases} 1 & f(x,y)\leq t \\ 0 & f(x,y)>t \end{cases}$$





(2) 阈值为灰度范围 $[t_1, t_2]$ ，阈值化为：

$$g(x,y) = \begin{cases} 1 & t_1 \leq f(x,y) \leq t_2 \\ 0 & \text{其它} \end{cases}$$

(3) 设阈值为 t ，半阈值化为：

$$g(x,y) = \begin{cases} f(x,y) & f(x,y) \geq t \\ 0 & \text{其它} \end{cases}$$



(4) 图像阈值分割的通式为：

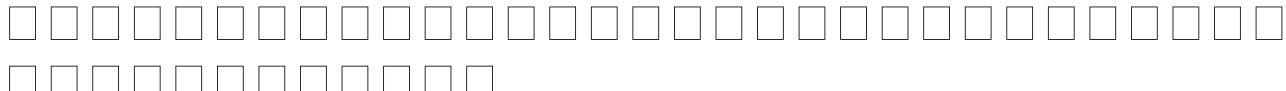
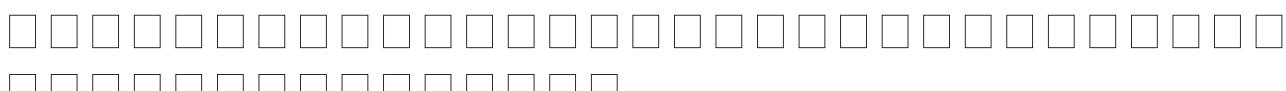
$$g(x,y) = \begin{cases} Z_E & f(x,y) \in Z \\ Z_B & \text{其它} \end{cases}$$

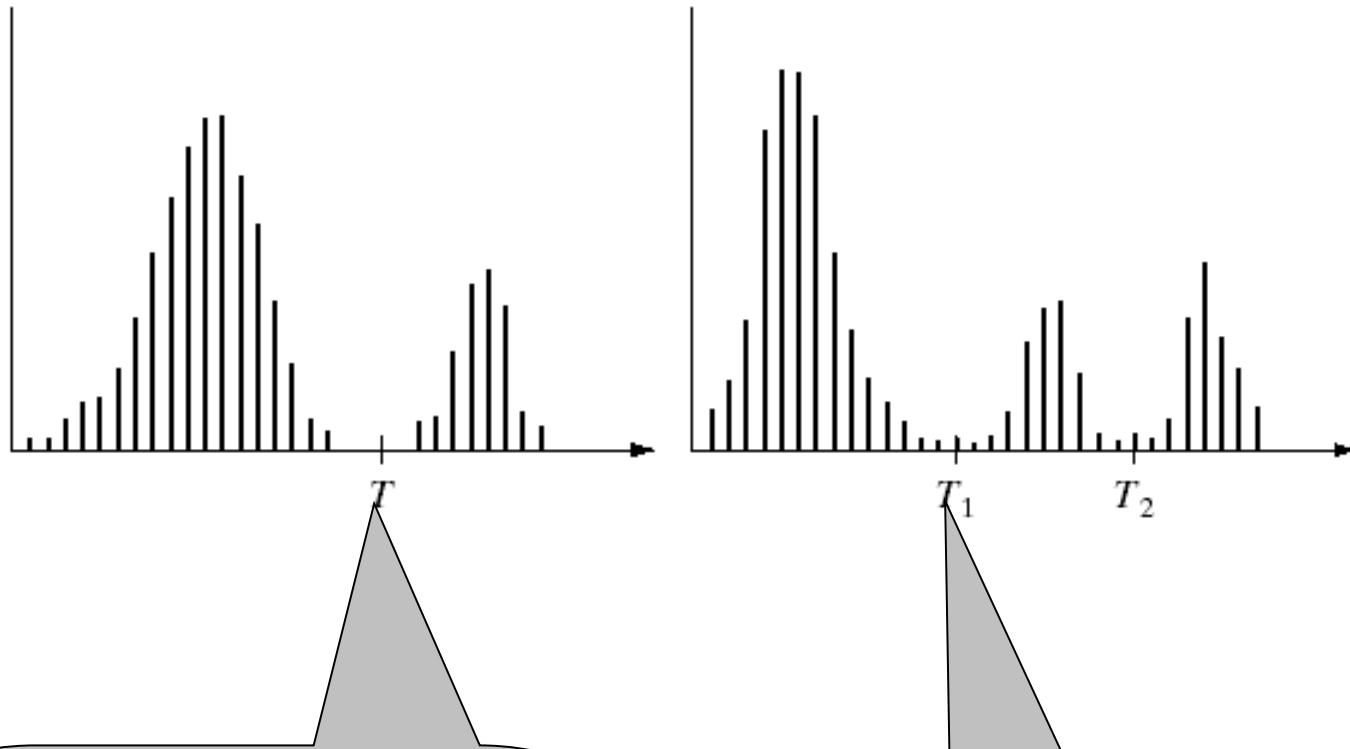
—— 阈值

可见，阈值化分割算法主要有两个步骤：

- 1) 确定合适的分割阈值
- 2) 将每个像素值与分割阈值比较以划分像素归属。

2、阈值的分类

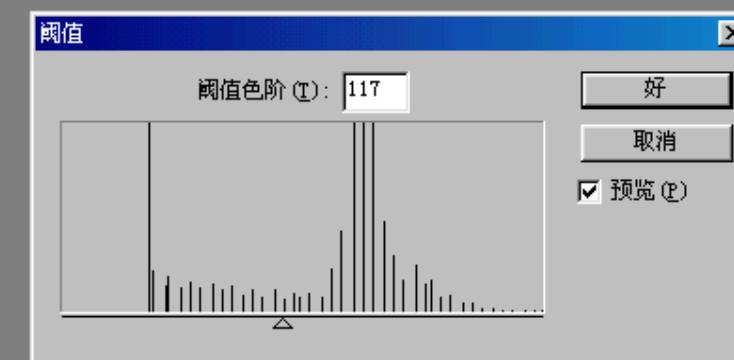
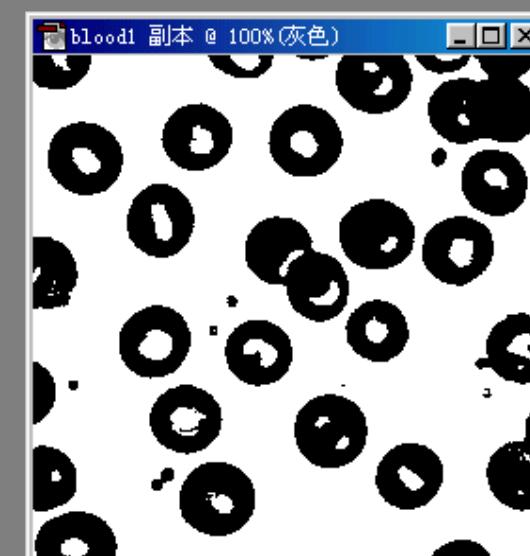
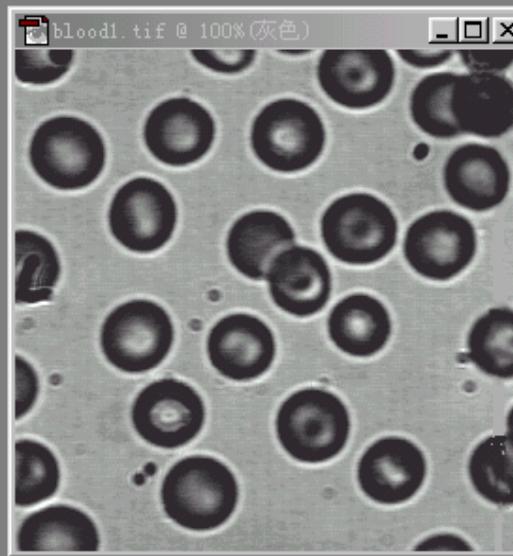
阈值类型	适用情况
全局阈值	图像中目标和背景有明显差别，而且这种差别在图像中的任何位置都几乎相同。
局部阈值	目标和背景的对比度随着位置的不同可能有较大变化（比如光照不均）。
单阈值	图像中只有单个目标和单个背景的情况，或者只对图像进行简单的划分。
多阈值	
直接阈值	
间接阈值	



采用一个 T 就可以
将两个模式分离开

需要多个 T 分割

取样大小: 取样点



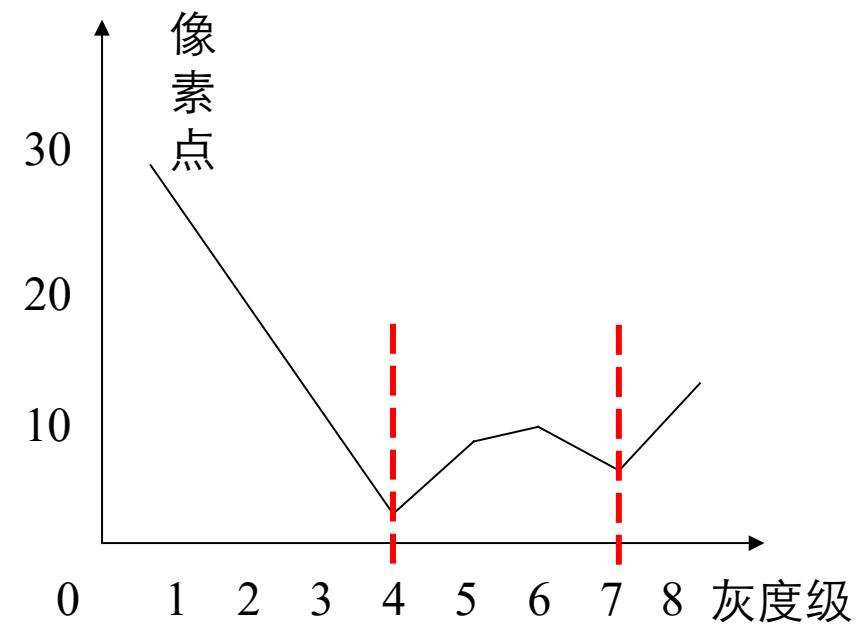


3、阈值分割的特点

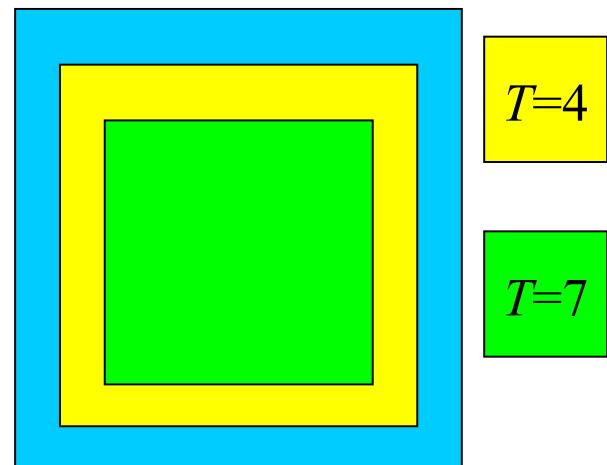
- 1) 对于物体与背景有较强对比的图像分割特别有效；
- 2) 计算简单；
- 3) 总能用封闭而且连通的边界定义不交叠的区域；
- 4) 可以推广到非灰度特征，如果物体同背景的区别不在灰度值，而是其它特征，如纹理、颜色等，可以先计算那种特征，再转化为灰度图，然后就可以利用阈值分割技术。

1	1	1	1	1	1	1	1
1	4	5	6	5	6	5	1
1	5	7	8	8	8	6	1
1	6	8	8	8	7	6	1
1	5	7	8	8	8	6	1
1	4	8	7	7	7	6	1
1	6	5	4	5	6	5	1
1	1	1	1	1	1	1	1

(a) 图像



(b) 直方图



(c) 分割结果



6.2.2 阈值选取方法

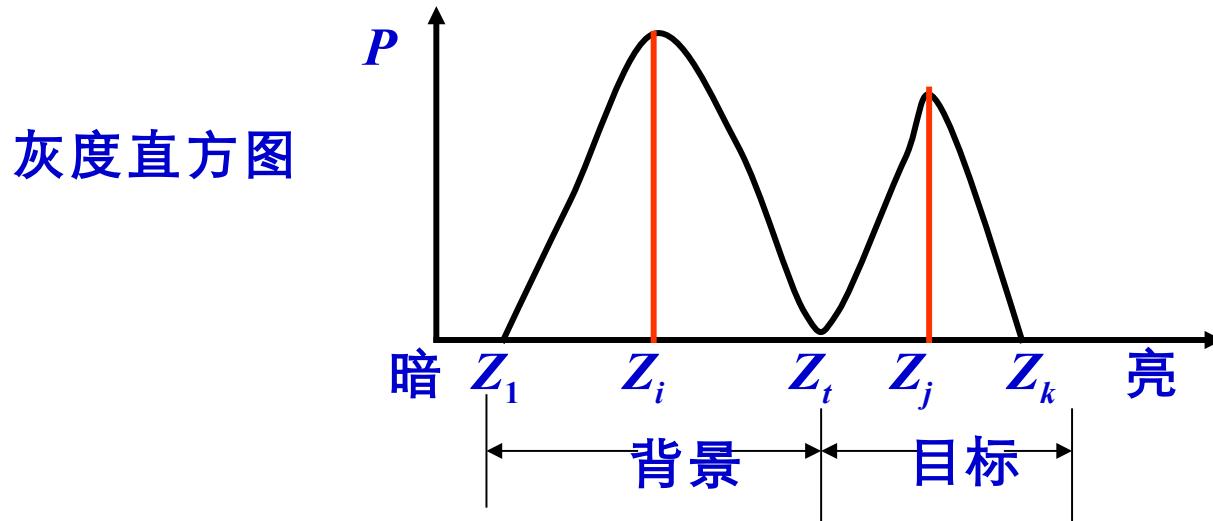
- 1、根据直方图谷点确定阈值
- 2、最优阈值（最小错误概率阈值）
- 3、最大类间方差阈值
- 4、 P 参数法阈值



1、根据直方图谷点确定阈值

1) 基本思想

目标或背景内部的像素灰度高度相关，而在它们交界处两边的像素灰度则有很大的差别，图像的直方图基本上可以看作是由目标和背景两个单峰直方图叠加而成。若目标和背景灰度差别较大时，图像的直方图应该是双峰的。





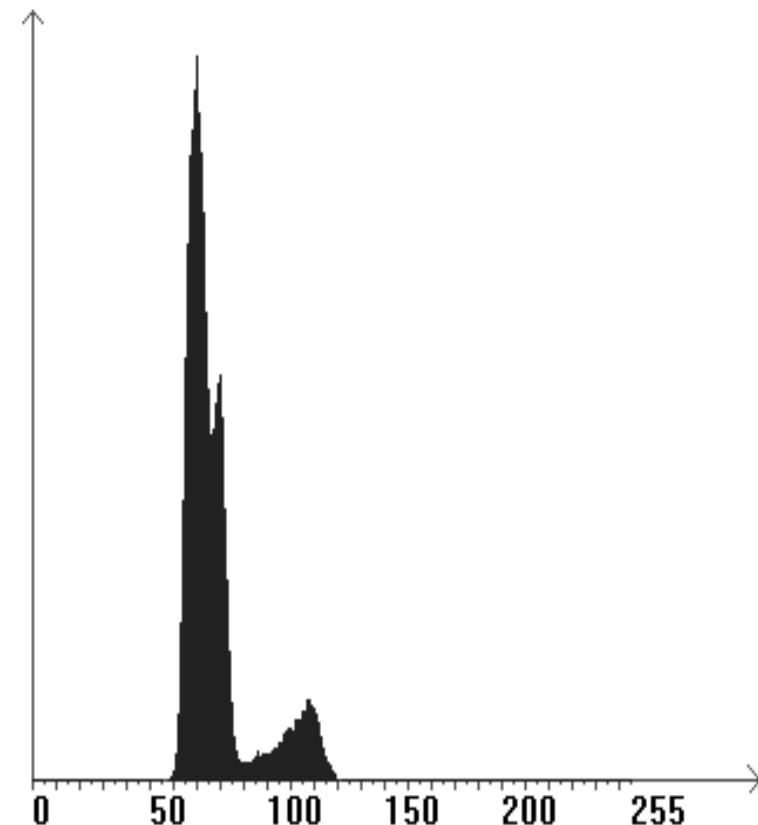
2) 阈值取法：取直方图谷底（最小值）的灰度值为阈值 T 。



(a) 原始图像



(c) 二值化图像

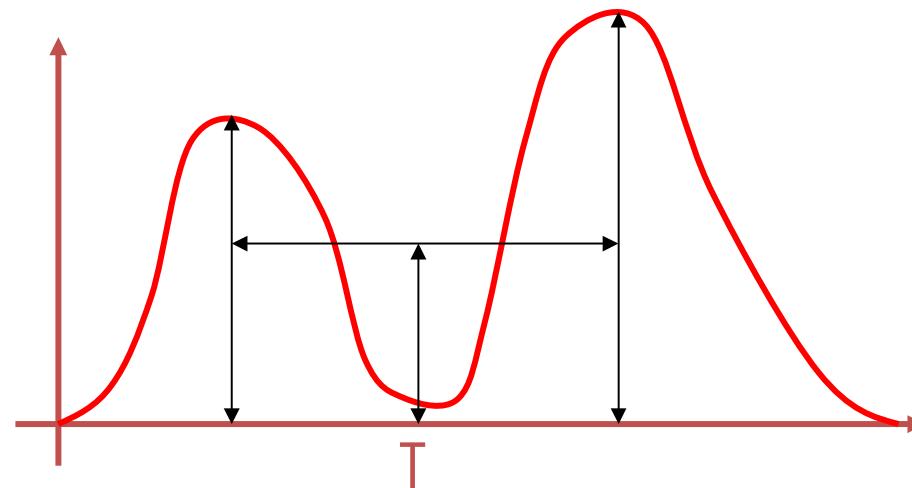


(b) 原图的直方图



3) 存在的缺点：会受到噪声的干扰，最小值不是预期的阈值，而偏离期望的值。

改进1：取两个峰值之间某个固定位置，如中间位置上。由于峰值代表的是区域内外的典型值，一般情况下，比选谷底更可靠，可排除噪声的干扰。

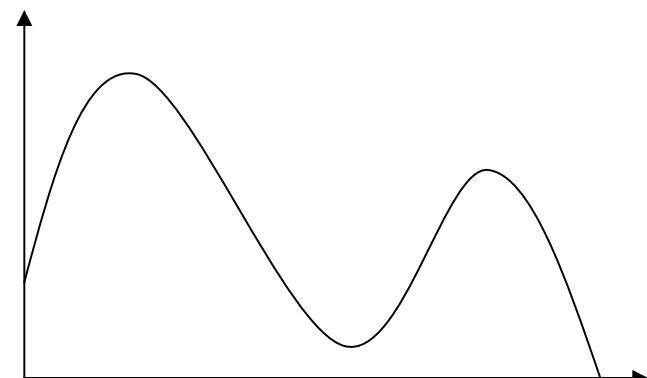


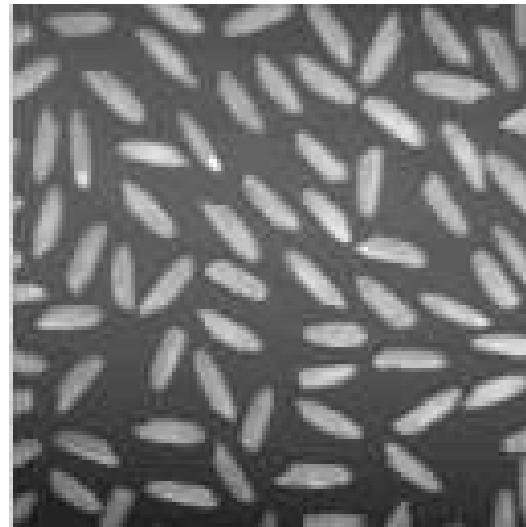


改进 2：加强对噪声的处理。对直方图进行平滑处理。

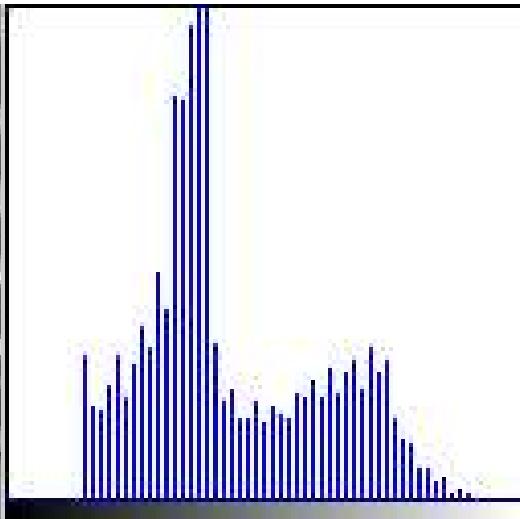
$$\frac{\partial h(z)}{\partial z} = 0, \quad \frac{\partial^2 h(z)}{\partial z^2} > 0$$

其中， $h(z)$ ——直方图

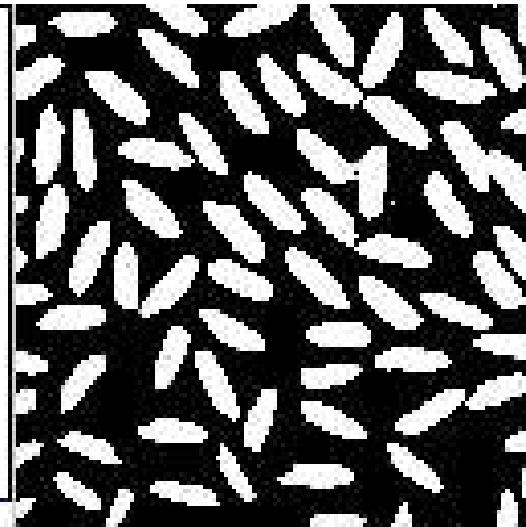




(a) 原图
值化结果



(b) 直方图



(c) 二

rice 图像双峰法分割 ($T=120$)



2、最优阈值（最小错误概率阈值）

1) 问题的提出：

- a: 图像中目标和背景的灰度分布过于分散或者它们的分布有部分交错，这使得图像的直方图双峰的方差太大，均值相距太近，而没有明显的双峰 - 谷底现象，很难找出合适的阈值。
- b: 更有甚者，即使直方图有较为明显的双峰一谷，但因为它是两个单峰直方图的叠加，很有可能使得此时的谷点并不是最准确的阈值点。

为此可以利用一些统计学知识，在某种评价的标准下寻优，找出相应意义下的最优阈值。



2) 最优阈值的思想

分割的结果在错分概率准则下达到最优，因此称之为最优阈值或最小错误概率阈值。

3) 具体做法

- (1) 设定目标物和背景的概率及其灰度分布概率密度函数
- (2) 给定一个阈值 t 下，求每类的分割错误概率
- (3) 求此阈值下总分割错误概率 $e(t)$
- (4) 由总分割错误概率 $e(t)$ 的极小值求解最优阈值 T

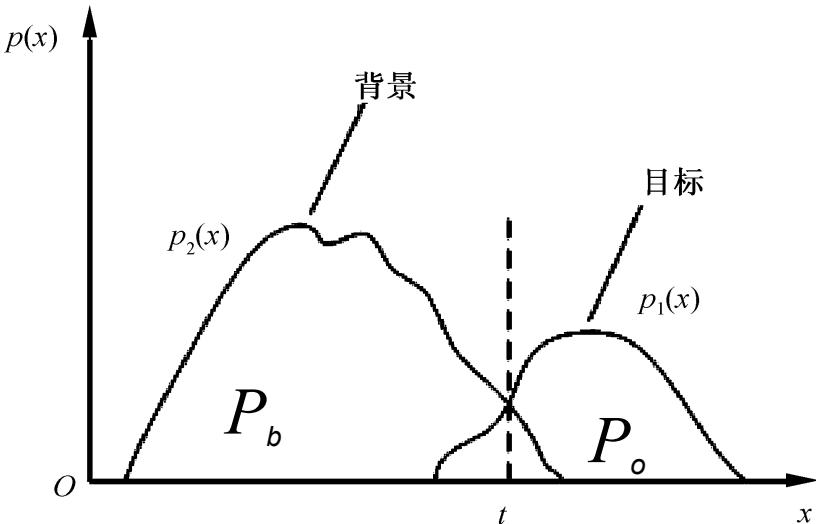


假设目标和背景的灰度概率分布如下
图：

分割错误 1 : $e_1 = \int_t^{\infty} p_1(x)dx$

分割错误 2 : $e_2 = \int_{-\infty}^t p_2(x)dx$

分割总错误概率：



$$e(t) = P_o e_1 + P_b e_2 = P_o \int_t^{\infty} p_1(x)dx + P_b \int_{-\infty}^t p_2(x)dx$$

$$\frac{de(t)}{dt} = 0$$

$$P_o p_1(T) = P_b p_2(T)$$

最优阈值
应满足的
条件

举例 若 $p_1(x)$ 和 $p_2(x)$ 为正态(高斯)分布, 即

:

$$p_1(x) = N(\mu_1, \sigma_1^2) \quad p_2(x) = N(\mu_2, \sigma_2^2)$$

则: $P_1(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp[-(x - \mu_1)^2 / 2\sigma_1^2]$

$$P_2(x) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp[-(x - \mu_2)^2 / 2\sigma_2^2]$$

将这两式代入 $P_o P_1(T) = P_b P_2(T)$ 中, 化简后可得:

$$AT^2 + BT + C = 0$$

$$A = \sigma_1^2 - \sigma_2^2, \quad B = 2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2),$$

$$C = \sigma_1^2\mu_2^2 - \sigma_2^2\mu_1^2 + 2\sigma_1^2\sigma_2^2 \ln(\sigma_2 P_o / \sigma_1 P_b)$$



总误差最小时一般有两个解。如果两个区域的方差相等，则只有一个最优阈值（ T ），即：

if $\sigma^2 = \sigma_1^2 = \sigma_2^2$,

then $T = \frac{\mu_1 + \mu_2}{2} + \frac{\sigma^2}{\mu_1 - \mu_2} \ln(P_b / P_0)$



if $P_b = P_o = 1/2$

then $T = (\mu_1 + \mu_2) / 2$



数字图像处理

(Digital Image Processing)

山东科技大学信息与电气工程学院

范 迪



6.2 阈值分割法

- 1、根据直方图谷点确定阈值
- 2、最优阈值
- 3、最大类间方差阈值
- 4、 P 参数法阈值

6.3 边缘检测分割 法

- 6.3.1 边缘分割法原理
- 6.3.2 边缘检测算子与分割
- 6.3.3 模板匹配法边缘检测
- 6.3.4 曲面拟合法边缘检测



3、最大类间方差阈值

1) 基本思想

认为图像中目标或背景内的像素灰度相近，而它们之间往往存在较大的差异。选择一个阈值，基于它的分割能使得目标和背景区域之间的总体差别最大，在某种程度上，可以认为这样的分割结果已达到了最优。区域间的这种差别常用方差来描述。

2) 具体做法

图像 $f(x,y)$ 具有 L 级灰度，灰度分布为 p_i ，假设以阈值 k 来分割为低灰度的目标 R_1 和高灰度的背景 R_2 ，则：

关于 R_1

概率： $P_1 = \sum_{i=0}^{k-1} p_i = p(k)$

平均灰度： $\mu_1 = \frac{1}{P_1} \sum_{i=0}^{k-1} i p_i = \frac{\mu(k)}{p(k)}$

方差： $\sigma_1^2 = \sum_{i=0}^{k-1} (i - \mu_1)^2 p_i / P_1$

关于 R_2

$$P_2 = \sum_{i=k}^{L-1} p_i = 1 - p(k)$$

$$\mu_2 = \frac{1}{P_2} \sum_{i=k}^{L-1} i p_i = \frac{\mu - \mu(k)}{1 - p(k)}$$

$$\sigma_2^2 = \sum_{i=k}^{L-1} (i - \mu_2)^2 p_i / P_2$$

其中： $\mu = \sum_{i=0}^{L-1} i p_i = \sum_{i=0}^{k-1} i p_i + \sum_{i=k}^{L-1} i p_i$ 或 $P_1 \mu_1 + P_2 \mu_2 = \mu$ —— 图像的平均灰度



R_1 、 R_2 的类内方差定义为 $\sigma_{k1}^2 = P_1\sigma_1^2 + P_2\sigma_2^2$

两个区域类间方差定义为：

$$\begin{aligned}\sigma_k^2 &= P_1(\mu_1 - \mu)^2 + P_2(\mu_2 - \mu)^2 \\&= p(k)\left[\frac{\mu(k)}{p(k)} - \mu\right]^2 + [1 - p(k)]\left[\frac{\mu - \mu(k)}{1 - p(k)} - \mu\right]^2 \\&= \frac{[\mu(k) - \mu p(k)]^2}{p(k)} + \frac{[\mu p(k) - \mu(k)]^2}{1 - p(k)} \\&= \frac{[\mu p(k) - \mu(k)]^2}{p(k)[1 - p(k)]}\end{aligned}$$

总体方差定义为： $\sigma^2 = \sigma_{k1}^2 + \sigma_k^2$

引入关于 k 的等价判决准则：

$$\lambda(k) = \frac{\sigma_k^2}{\sigma_{k1}^2}$$

类间 / 类内

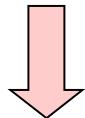
$$\eta(k) = \frac{\sigma_k^2}{\sigma^2}$$

类间 / 总方差

$$\kappa(k) = \frac{\sigma^2}{\sigma_{k1}^2}$$

总方差 / 类内

三个准则是等效的，把使 R_1, R_2 两类得到最佳分离的 k 值作为最佳阈值，将 $\lambda(k)$ 、 $\eta(k)$ 、 $\kappa(k)$ 定义为最大判决准则。



由于 σ_{k1}^2 是基于二阶统计特性，而 σ_k^2 是基于一阶统计特性，它们都是阈值 k 的函数，而 σ^2 与 k 值无关，因此三个准则中 $\eta(k)$ 最为简单，因此选其作为准则，可得到最佳阈值 k^*

$$k^* = \arg \max_{0 \leq k \leq L-1} \eta(k)$$



3) 特点

显然，方差是阈值 k 的函数， k 发生变化， R_1 和 R_2 两个区域之间的方差也随着发生变换，选择 K 值使得方差取得极大值，此时的 K 就是能够使得分割结果达到最大方差的阈值，称之为最大类间方差阈值。

最大类间方差阈值的求解过程不需要人为设定其他参数，完全由计算机自动选取。它不仅适用于两个区域的分割，也可以推广到多个区域的情形。



4、P参数法阈值

1) 基本设计思想

对固定分辨率下的目标物，根据目标物在画面中所占的比例来选择阈值，进行二值化处理。

它对于已知目标物在画面中所占比例情况下的图像分割比较有效。



2) 具体做法

假设目标物为暗，背景为亮；

先试探性地给出一个阈值 1，统计目标物的像素点数在整幅图中所占的比例是否满足要求，是则阈值合适；否则，阈值则偏大（右）或者偏小（左），再进行调整，直到满足要求。





P参数法阈值的算法：

1. 计算得到原图的灰度直方图 h ；
2. 输入目标物所占画面的比例 p ；
3. 尝试性地给定一个阈值 $T_h = T_{h0}$ ；
4. 计算在 T_h 下判定的目标物的像素点数 $N \sum_{k=0}^{T_h} h(k)$
5. 设图像的大小为 $m * n$ ，判断 $p_s = N / (m * n)$ 是否接近 p ？

是，则输出结果，否则， $T_h = T_h + dT$; (if

$p_s < p$, 则

$dT > 0$; else $dT < 0$) , 转 4 , 直到



6.3 边缘检测分割法

6.3.1 边缘分割法原理

6.3.2 边缘检测算子与分割

6.3.3 模板匹配法边缘检测

6.3.4 曲面拟合法边缘检测



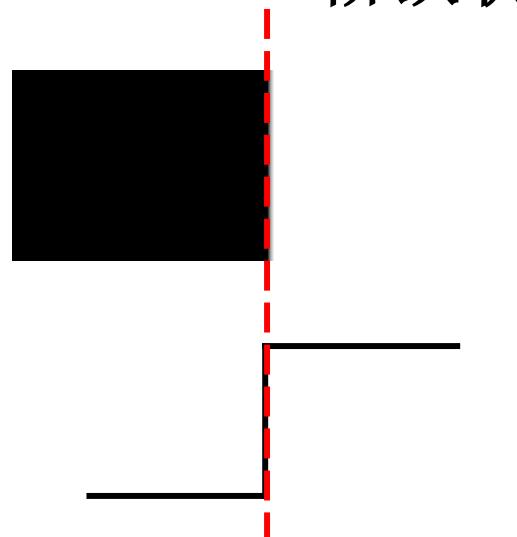
6.3.1 边缘边缘分割法原理

1、边缘的定义

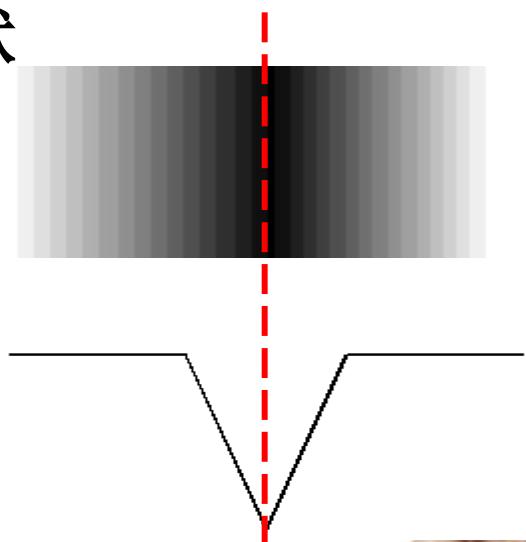
图像中像素灰度有阶跃变化或屋顶变化的那些像素的集合。

2、边缘的分类

图像：



屋顶状

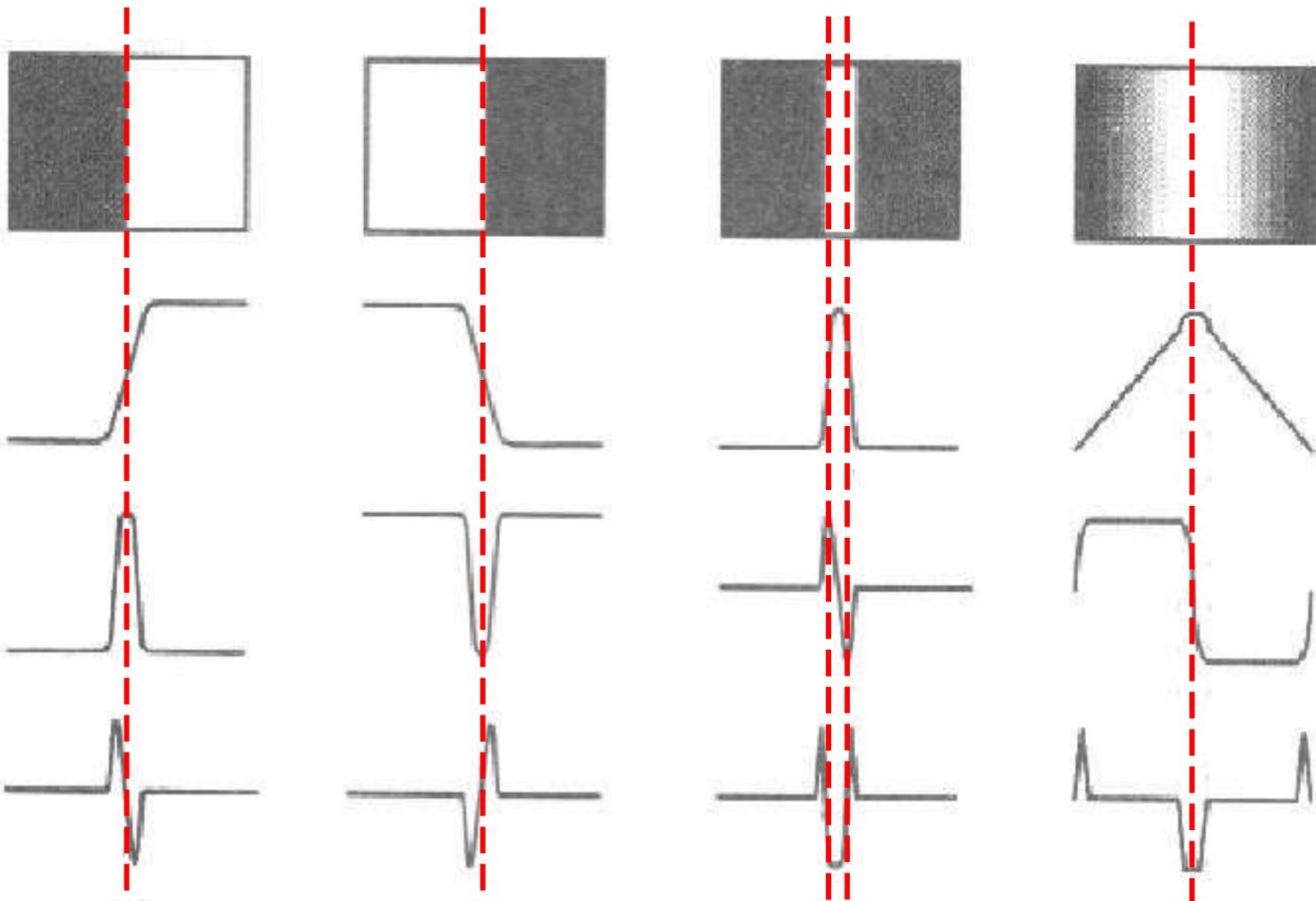


剖面：



各种边缘其一阶、二阶导数特点

图像：



剖面：

一阶
导数：

二阶
导数：



数字图像处理

(Digital Image Processing)

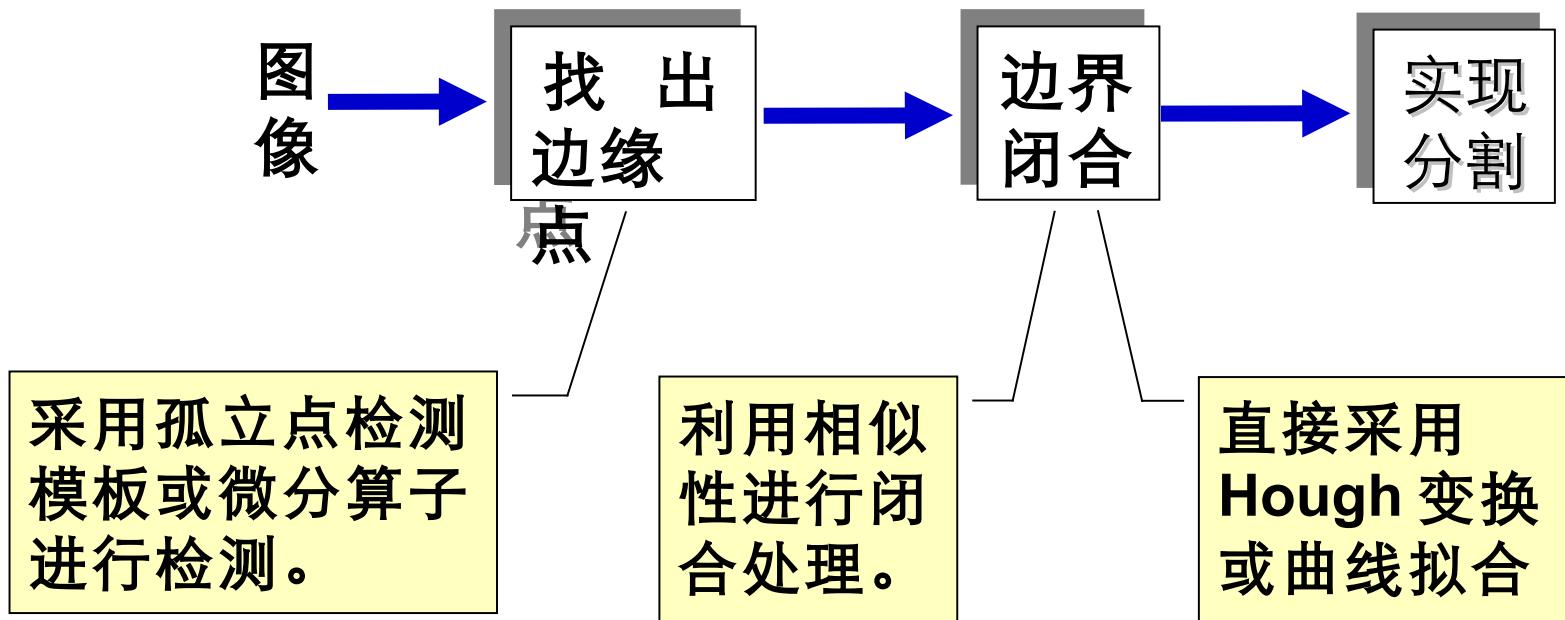
山东科技大学信息与电气工程学院

范 迪



3、边缘分割法原理

其基本思想是通过寻找图像中不同区域的边界，从而实现图像分割。它是基于边界的一大类图像分割方法。





6.3.2 边缘检测算子与分割

1、边缘检测算子 主要有一阶、二阶局部微分算子

•				
□ □ □ □	□ □ □ □ □ □	Roberts □ □	Prewitt □ □	Sobel □ □
f_x		$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$
f_y	$\begin{bmatrix} -1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$

□ □ □ □ □ □			LOG/Marr □ □
1 -2 1 -2 4 -2 1 -2 1	0 -1 0 -1 4 -1 0 -1 0	-1 -1 -1 -1 8 -1 -1 -1 -1	$\begin{bmatrix} 1 & 0 & 1 \\ 0 & -4 & 0 \\ 1 & 0 & 1 \end{bmatrix}$



2、微分算子检测边缘的原理

1) 一阶微分算子检测边缘的基本思想:

检测图像一阶导数的峰值或者谷值确定边缘，可用一阶微分算子和图像卷积实现。

2) 二阶微分算子检测边缘的基本思想:

检测图像二阶导数的零点确定边缘，可用二阶微分算子和图像卷积实现，并可通过二阶导数的正负判断像素在明区还是暗区。



3、一阶微分算子检测边缘

1) 梯度算子

函数 $f(x,y)$ 在 (x,y) 处的梯度向量为 $\nabla f = [\partial f_x \quad \partial f_y] = [\partial f / \partial x \quad \partial f / \partial y]$

梯度向量的大小为 : $G[f(x,y)] = \sqrt{[\partial f / \partial x]^2 + [\partial f / \partial y]^2}$

近似为 : $G[f(x,y)] = \max\{|\partial f_x|, |\partial f_y|\}$ 或 $G[f(x,y)] = |\partial f / \partial x| + |\partial f / \partial y|$

梯度向量的方向角为 $\Phi(x,y) = \arctg^{-1}[\partial f_y / \partial f_x]$



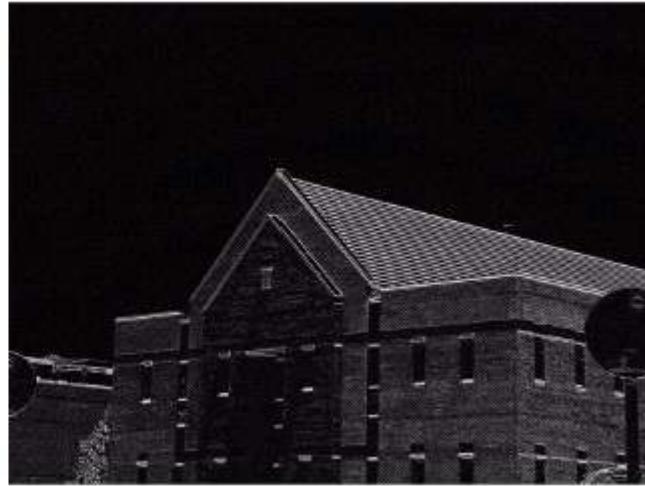
为了检测边缘点，选取适当的阈值 T ，对梯度图像进行二值化，则有：

$$g(x, y) = \begin{cases} 1 & G[f(x, y)] \geq T \\ 0 & \text{else} \end{cases}$$

- a. 直接计算 ∂f_x ∂f_y 可以检测到边的存在。
- b. $|\partial f_x|$ 对沿着 x 轴方向的边产生的响应最强；
 $|\partial f_y|$ 对沿着 y 轴方向的边产生的响应最强。
- c. 有的算子在微分的同时，具有平滑作用，能够滤除噪声。



(a) 原图



(b) x 方向上的梯度分量



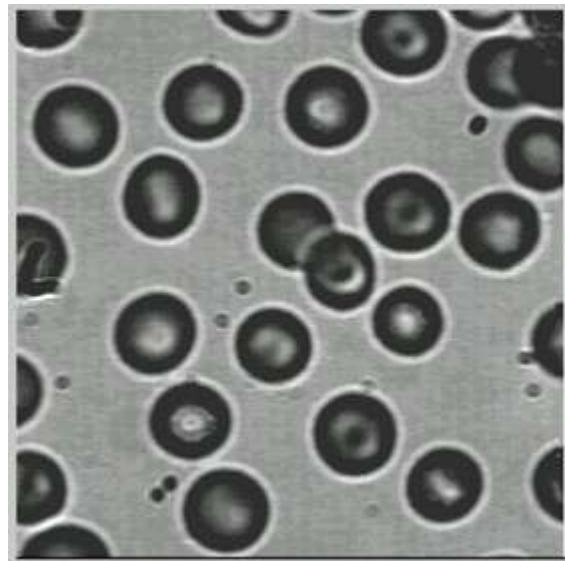
(c) y 方向上的梯度分量



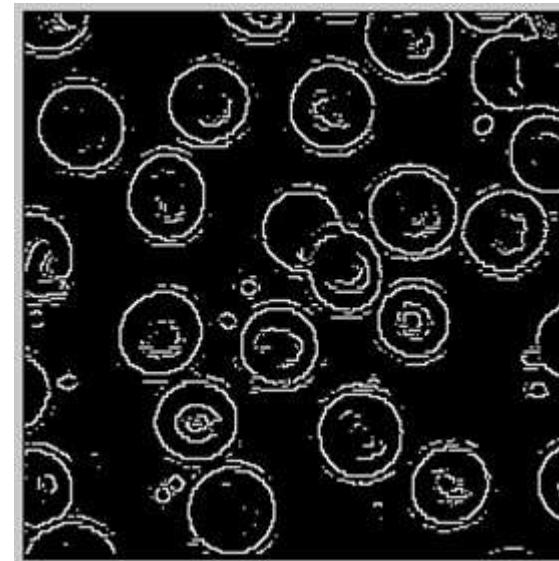
(d) 梯度图像



边缘检测分割的例子

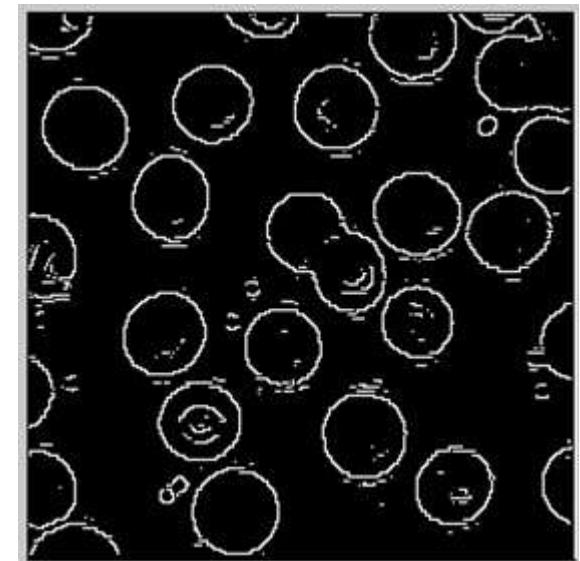


原图
 $T=0.12$ (Roberts)



$T=0.08$ (Roberts)

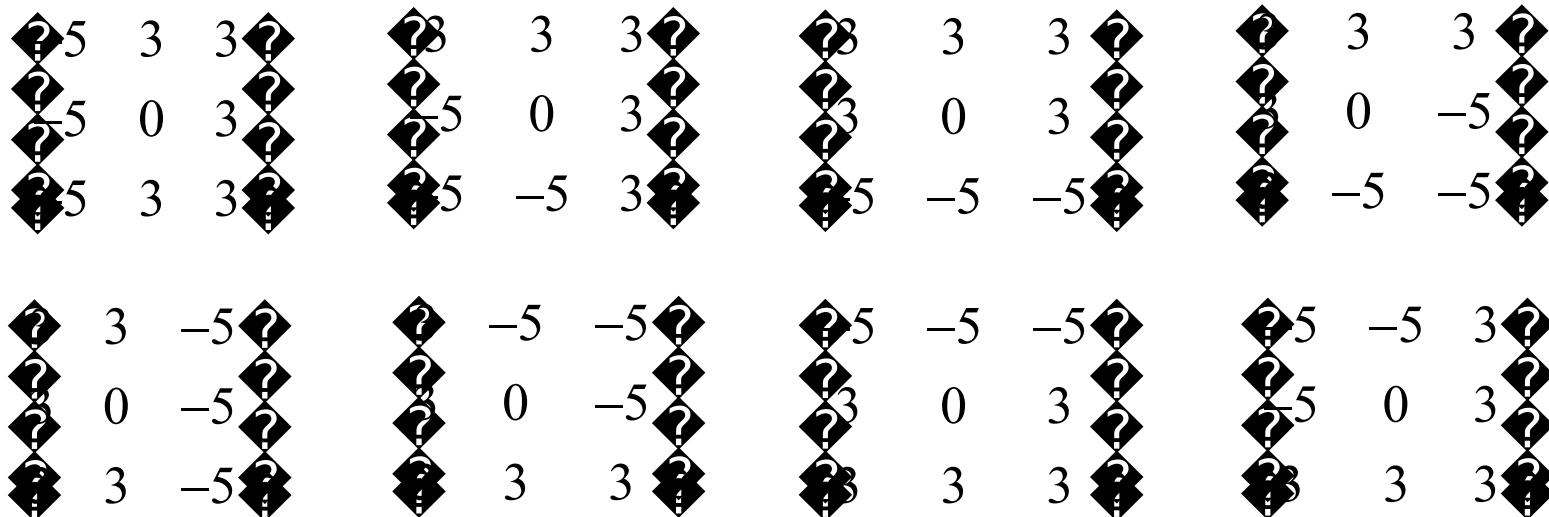
$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$



2) 其它方向算子

Roberts , Prewitt , Sobel 等算子，只包含两个方向的模板，可检测边缘方向少。方向模板方向越多就能检测更多方向的边缘。

(1) Kirsch 方向模板：8个方向依次成 45° 夹角





使用方法： 8个模板分别与图像卷积，检测8个方向的灰度变化，变化最大的方向即是边缘的方向。

Kirsch 算子的方向模板也可以有更大的尺寸， 5×5 的 Kirsch 算子的前 4 个方向模板如下：

$$\begin{array}{cccccc} \begin{matrix} 1 & -1 & 0 & 1 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{matrix} & \begin{matrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{matrix} & \begin{matrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{matrix} & \begin{matrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{matrix} \\ \begin{matrix} 1 & -1 & 0 & 1 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{matrix} & \begin{matrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{matrix} & \begin{matrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{matrix} & \begin{matrix} 1 & 0 & -1 & -1 & -1 \\ 0 & 1 & -1 & -1 & -1 \\ -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & 1 \end{matrix} & \begin{matrix} 1 & 1 & 1 & 0 & ? \\ 1 & 1 & 0 & -1 & ? \\ 1 & 0 & -1 & -1 & ? \\ 0 & -1 & -1 & -1 & ? \\ -1 & -1 & -1 & -1 & ? \end{matrix} \end{array}$$



(2) Prewitt 对角模板

0	1	1
-1	0	1
-1	-1	0

-1	-1	0
-1	0	1
0	1	1

(3) Sobel 对角模板

0	1	2
-1	0	1
-2	-1	0

-2	-1	0
-1	0	1
0	1	2



(4) Canny 算子

常规边缘检测存在的问题：*a.* 位置不准； *b.* 非单像素宽等。

好的检测算子需满足三个指标：

- a) 高准确性：*** 多包含真边缘，少包含假边缘；
- b) 高精确度：*** 检测到的边缘应该在真正的边界上；
- c) 单像素宽：*** 选择性很高，只对边缘有唯一响应。

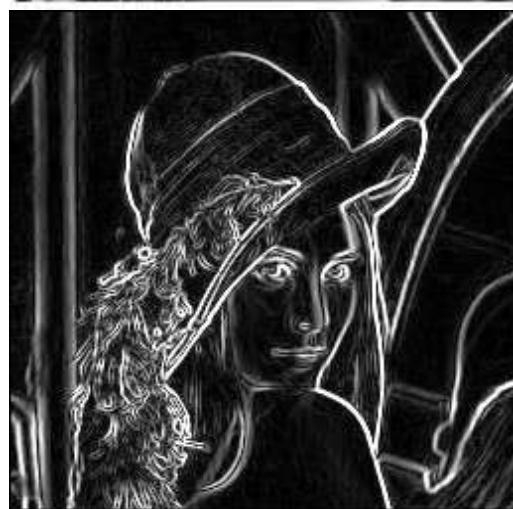
Canny 算子检测边缘的步骤

- ☛ 用高斯滤波模板平滑图像
- ☛ 计算平滑后图像梯度的幅值和方向
- ☛ 对梯度幅值应用非极大值抑制。遍历图像，若某个像素的梯度值与其梯度方向上前后两个像素的梯度值相比不是最大，那么这个像素不是边缘，其梯度置为 0。
- ☛ 用双阈值算法检测和连接边缘。使用两个阈值 T_1 和 T_2 ($T_1 > T_2$)，凡是大于 T_1 的一定是边缘，凡是小于 T_2 的一定不是边缘，如果检测结果介于 T_2 和 T_1 之间，则这个像素的邻接像素中有没有超过 T_1 的边缘像素，如果有，则该像素是边缘，否则就不是边缘。

(a) 原图



(b) Canny 算子



(c) Prewitt 算子



(d) Sobel 算子



4、二阶微分算子检测边缘

主要有拉普拉斯算子和 LOG 算子（也称 Marr 算子）。

（1）拉普拉斯算子

$$\text{二阶导数: } \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

1	-2	1
-2	4	-2
1	-2	1

-1	-1	-1
-1	8	-1
-1	-1	-1

-1	-1	-1
-1	8	-1
-1	-1	-1

拉普拉斯算子特点

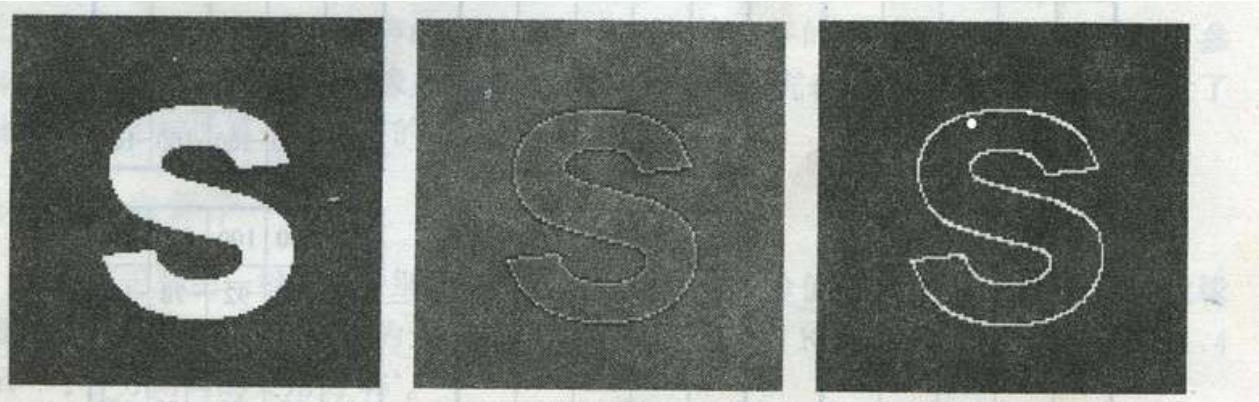
- a) 它是二阶导数，对噪声非常敏感。
- b) 拉普拉斯算子的幅值产生双边缘（最大负值和最大正值）

。

- c) 不能检测边缘的方向（无方向模板）。

在分割中所起的作用：

- a) 利用它的零交叉性质进行边缘预定位。
- b) 确定一个像素在边缘暗的一边还是亮的一边。





(2) LOG 算子

为克服二阶微分对噪声敏感的问题，先平滑图像，然后再运用拉氏算子检测边缘。

- a) 图像与高斯函数卷积进行滤波： $[h(x, y) * f(x, y)]$
- b) 对之拉氏算子卷积： $g(x, y) = \nabla^2 [h(x, y) * f(x, y)]$

$$g(x, y) = [\nabla^2 h(x, y)] * f(x, y) = \frac{1}{\pi \sigma^4} \left(\frac{r^2}{\sigma^2} - 2 \right) \exp\left(-\frac{r^2}{2\sigma^2}\right) * f(x, y)$$

其中： $r^2 = x^2 + y^2$

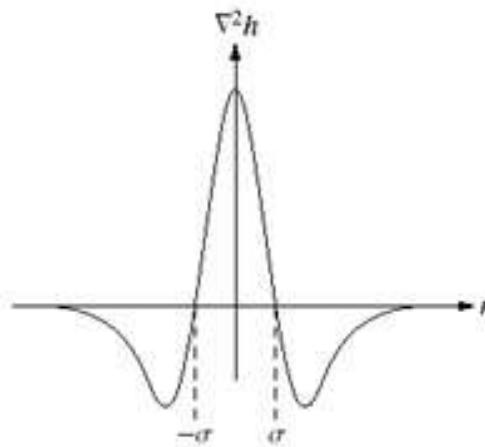
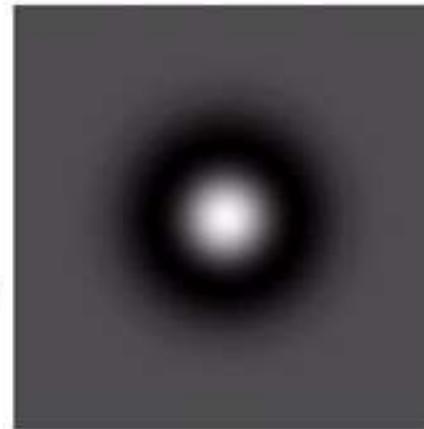
LOG 算子或 Marr 算子

LOG 函数的三维曲线、图像、剖面和模板

(a) 三维曲线



(b) 图像

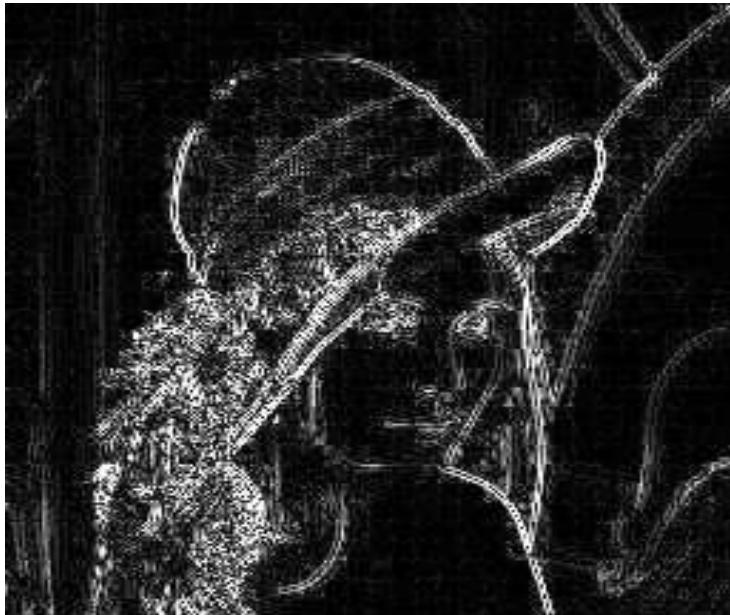


(c) 零交叉的剖面

0	0	-1	0	0
0	-1	-2	-1	0
-1	-2	16	-2	-1
0	-1	-2	-1	0
0	0	-1	0	0

(d) 近似的 5*5 模板 (e) 近似的 3*3 模板

$$\begin{array}{ccccc} \hat{e}1 & 0 & 1 & \bar{u} \\ \hat{e}0 & -4 & 0 & \bar{u} \\ \hat{e}1 & 0 & 1 & \bar{u} \end{array}$$



拉氏算子检测的边缘



LOG 算子检测的边缘



5、边缘接续和闭合

1) 边缘接续的目的

边缘接续的目的是要把间断的边连接起来形成封闭的边界。

2) 局部接续处理的原理

分析边缘检测结果中的每个点 (x,y) 的特性；在一个小的邻域（ 3×3 或 5×5 ）中把所有相似的点被连接，形成一个具有共同特性像素的边界。



6.3.3 模板匹配法边缘检测

1、基本思路

根据目标特征，建立**匹配模板**，通过考察模板和原图像中各子区域的**相似性**确定是否是目标。若相似，则认为该部分与模板相同，标记为目标。

2、相似性度量

差值测度法

$$L = \iint_D [f(x, y) - g(x, y)]^2 dx dy$$

$$L = \max_{(x, y) \in D} |f(x, y) - g(x, y)|$$

$$L = \iint_D |f(x, y) - g(x, y)| dx dy$$

相关测度法

$$L(k, l) = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{M-1} g(i, j) f(k+i, l+j)}{\sqrt{\sum_{i=0}^{M-1} \sum_{j=0}^{M-1} g^2(i, j) \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} f^2(k+i, l+j)}}$$



6.3.4 曲面拟合法边缘检测

1、基本思路

用一个平面或曲面去逼近图像中的局部灰度表面，然后用这个平面或曲面的梯度代替点的梯度，从而实现边缘检测，可减少噪声的干扰。

2、常用方法

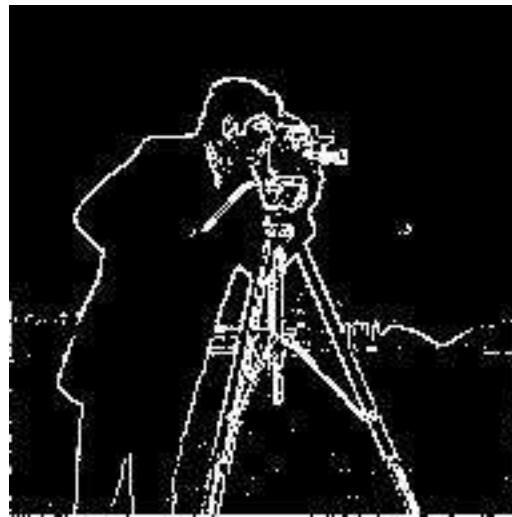
一次平面拟合 $f(x, y) = ax + by + c$

二次曲面拟合 $f(x, y) = ax^2 + bxy + cy^2 + dx + ey + g$

举例



原始图像



梯度算子



Roberts 算子



Prewitt 算子



Kirsch 算子



拉氏算子



曲面拟合法



数字图像处理

(Digital Image Processing)

山东科技大学信息与电气工程学院

范 迪



6.4 区域分割法

6.4.1 区域生长

6.4.2 分裂合并



6.4.1 区域生长

1、问题的引入

分割是将整幅图像区域 R 划分成 n 个子区域，它们要满足条件：

完备性

$$1) \bigcup_{i=1}^n R_i = R$$

独立性

$$2) R_i \cap R_j = \emptyset, \text{ 对所有的 } i \text{ 和 } j, i \neq j$$

单一性

$$3) P(R_i) = \text{TRUE}, i = 1, 2, \dots, n$$

互斥性

$$4) P(R_i \cup R_j) = \text{FALSE}, i \neq j$$

连通性

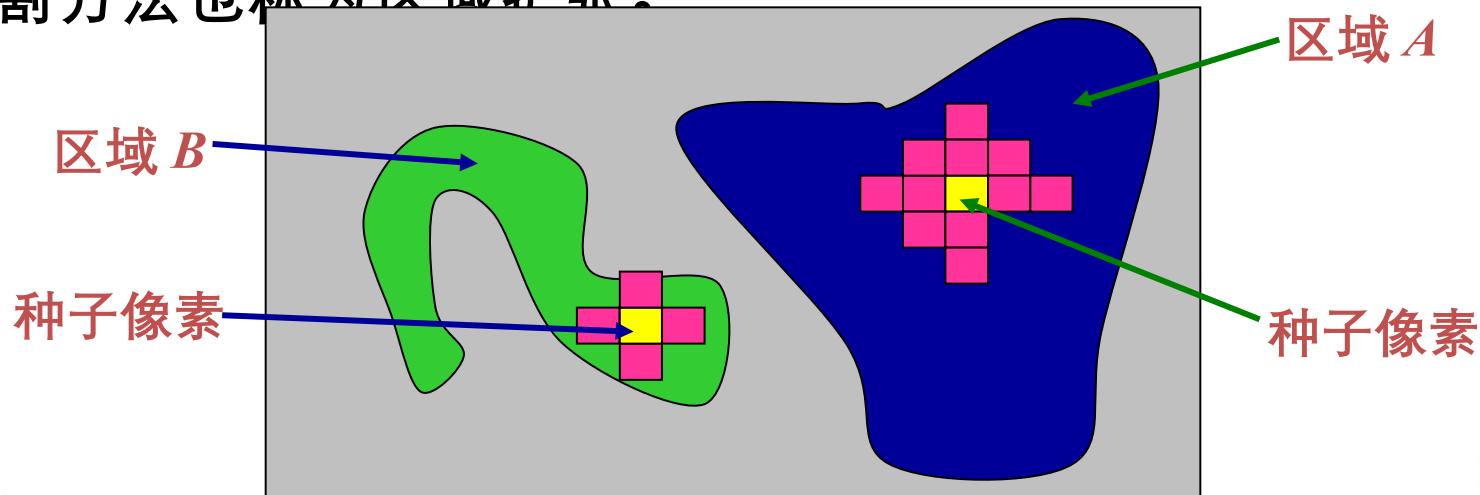
$$5) R_i \text{ 是一个连通的区域}, i = 1, 2, \dots, n$$

有的分割方法没有兼顾到全部条件。



2、基本思想

区域生长是先确定分成的区域个数及其特征，再分别为每个区域找一个具有代表性的种子，以种子为聚类中心不断地把图像中与之特性相同的相邻像素合并入该区域中，直到不能合并为止，最后形成特征不同的各个区域。这种分割方法也称为区域扩张。



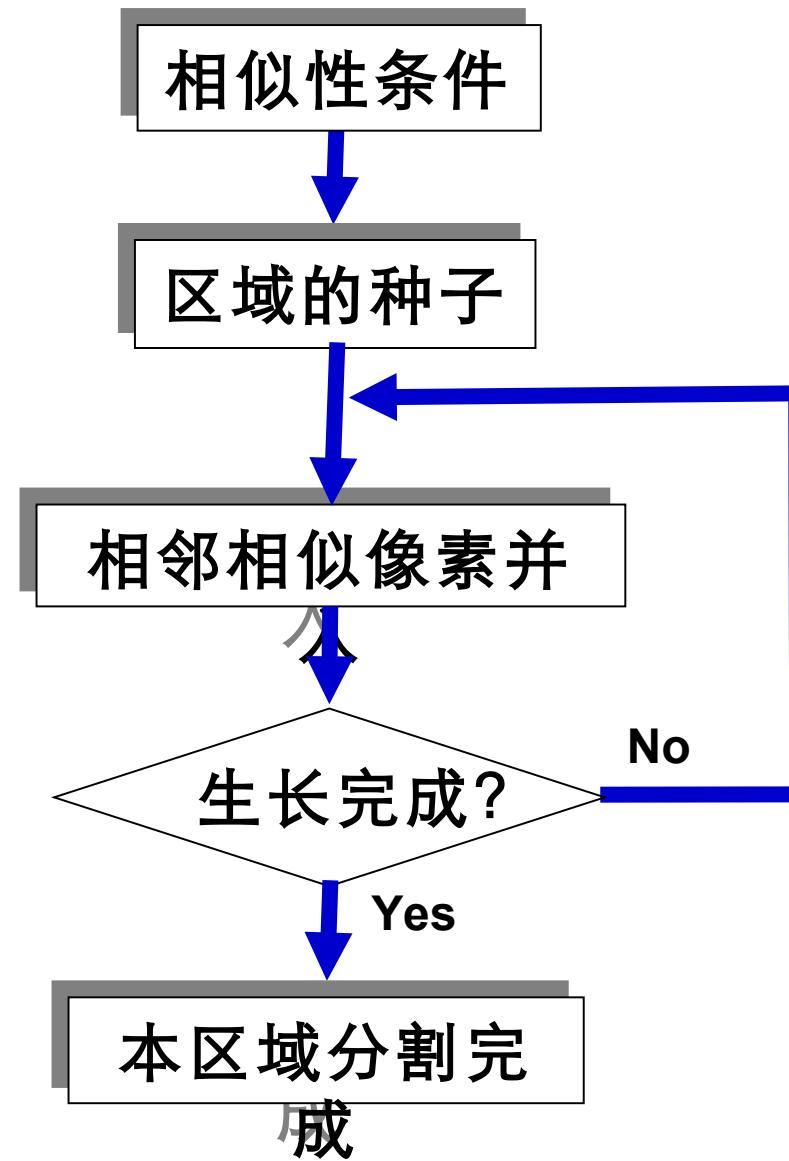


3、算法步骤

- 1) 先对要分割的区域找一个种子像素作为生长的起点；
- 2) 确定一个相似性标准（评判是否合并的条件）；
- 3) 按照相似性标准，将种子像素的邻域中具有与种子相同和相似性质的像素，合并到种子的区域中；
- 4) 将新像素当作新的种子像素继续进行上面的过程，直到再也没有满足条件的像素可被包括进来为止，这样就完成了一个区域的生长；
- 5) 截止其它区域的生长，直到将整幅图像完全分

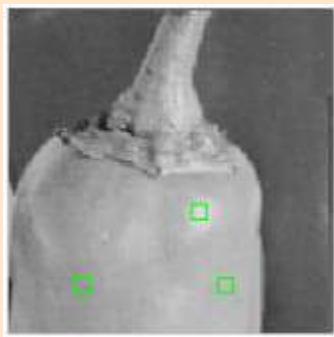


3、算法步骤





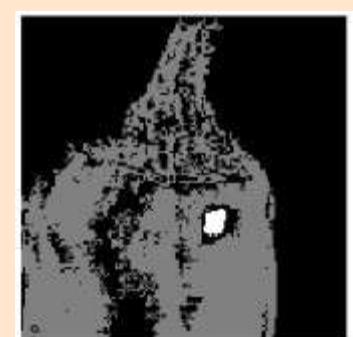
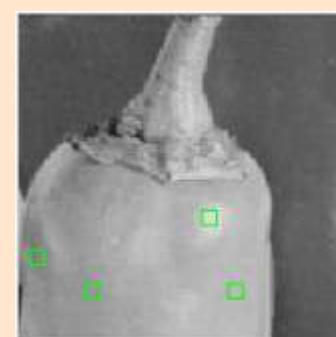
举例 1



原始图像及
种子点位置

三个种子点
区域生长结果

举例 2



原始图像及
种子点位置

四个种子点
区域生长结果



关键问题

- 种子点选择
- 相似性准则（生长准则）
确定
- 生长停止条件确定

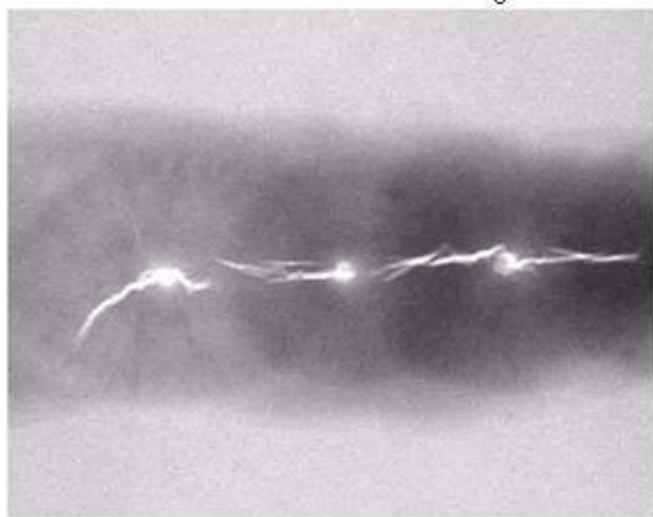
利用特征或先验知识

- 最亮的像素；
- 直方图极大值；
- 聚类中心处像素。

- 区域灰度差；
- 区域内灰度分布特性；
- 颜色、纹理、尺寸形状等

。

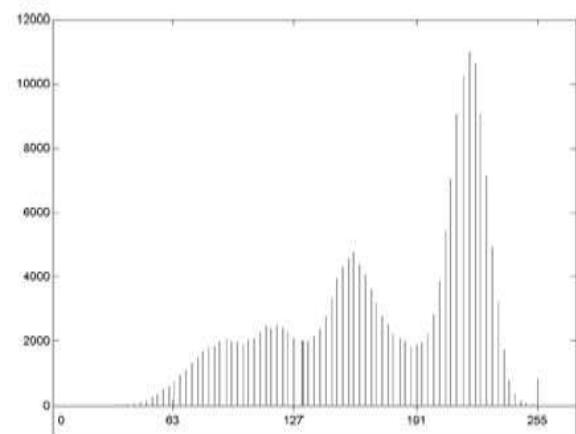
(a) 焊缝的 X 射线图



(b) 灰度为 255 的种子



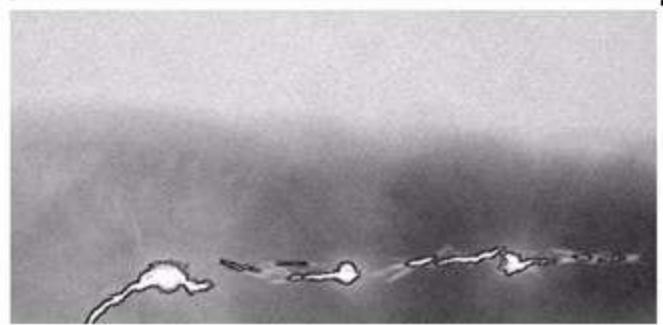
(e) 直方图



(c) 区域生长结果



(d) 分割后的边界



区域生长准则：任何像素和种子之间的灰度值绝对差小于 65。



6.4.2 分裂合并

1、问题的提出

区域生长受到种子、相似性准则等的影响很大。

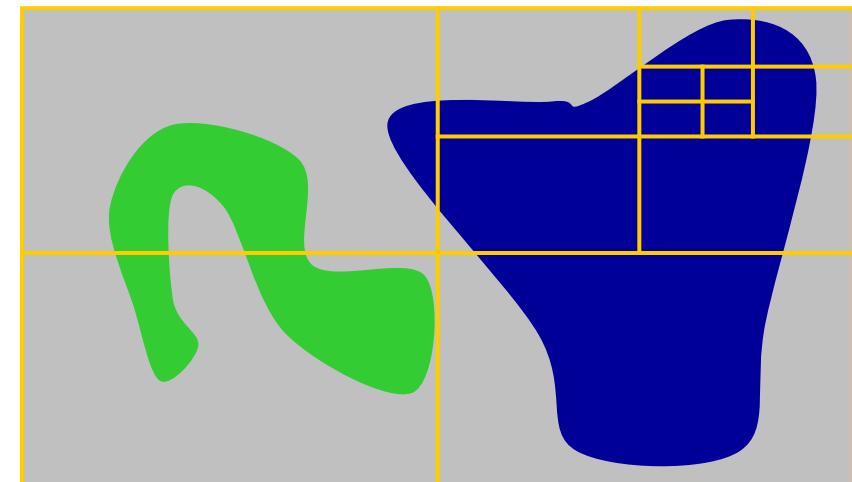
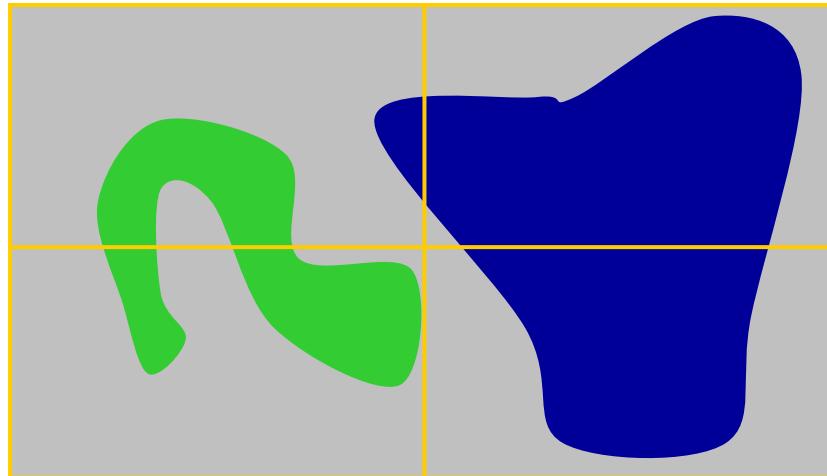
2、分裂合并的思想

从整幅图像开始，不断地逐级分裂，同时也把特性相同且相邻的区域合并，直到不能再分为止，最后得到各个子区域。



3、分裂 - 合并的做法

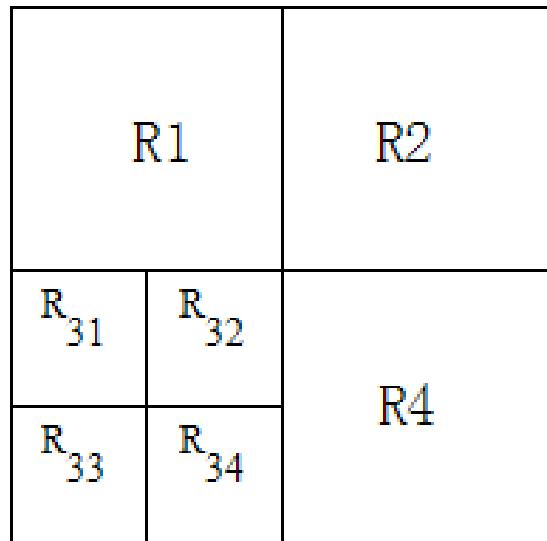
1) 分裂的做法：逐级四等分，直到要分裂的区域被分为单个像素为。



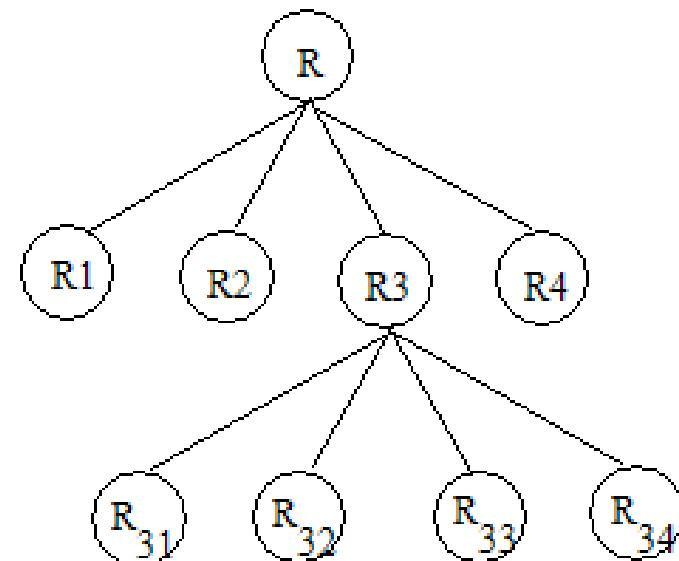
2) 合并的做法：把特性相同的相邻区域合并为一个区域。



分裂结果用四叉树表示



(a) 分裂图像

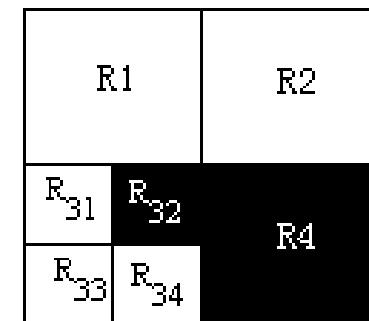
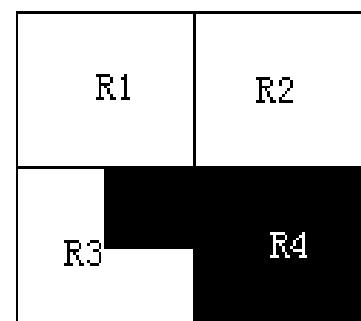
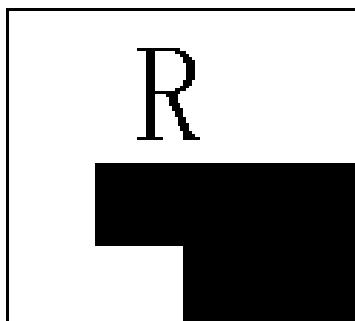


(b) 相应的四叉树结构



3) 分裂合并的做法

- a) 进行一次分裂；
- b) 考察所有相邻区域是否可以合并，若可以，则将其一一合并；
- c) 重复 a) b) 的分裂合并过程，直到分裂和合并都不能再进行为止。
注意：合并时一般先考虑同一个父节点下的 4 个区域，之后再扩展到其它父节点下同层次的区域。





数字图像处理

(Digital Image Processing)

山东科技大学信息与电气工程学院

范 迪



6.5 Hough 变换检测法

1、问题的提出

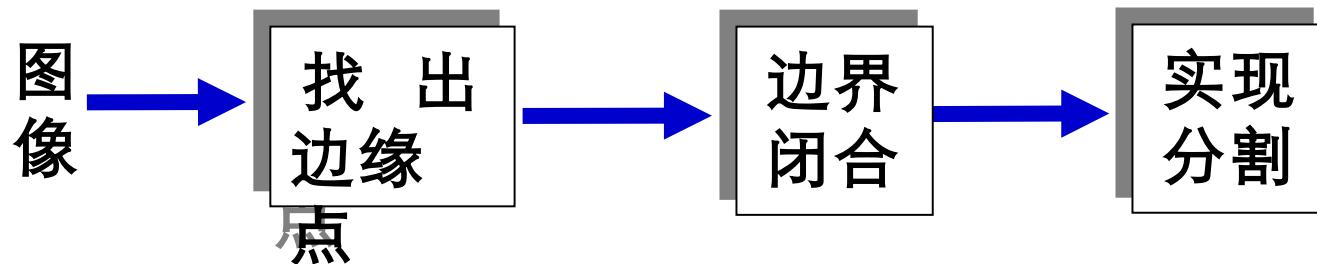
2、Hough 变换的基本原理

3、算法实现

4、Hough 变换的扩展



1、问题的提出



直接检测出封闭边界？

曲线拟合边界

Hough 变换



2、Hough 变换的基本原理

基本思想：利用一个空间和另一个空间的对偶关系，把原空间中的问题转换到它的对偶空间去求解，在对偶空间里问题变得相对简单。

在图像空间 xy 里，一条直线方程可表示为：

The diagram illustrates the Hough transform mapping. On the left, a blue curved arrow points from the equation $y = px + q$ to the equation $q = y - px$. The right side shows two parallel lines: the top line is labeled p — 直线斜率 (line slope) and the bottom line is labeled q — 直线截距 (line intercept).

$$y = px + q$$
$$q = y - px$$

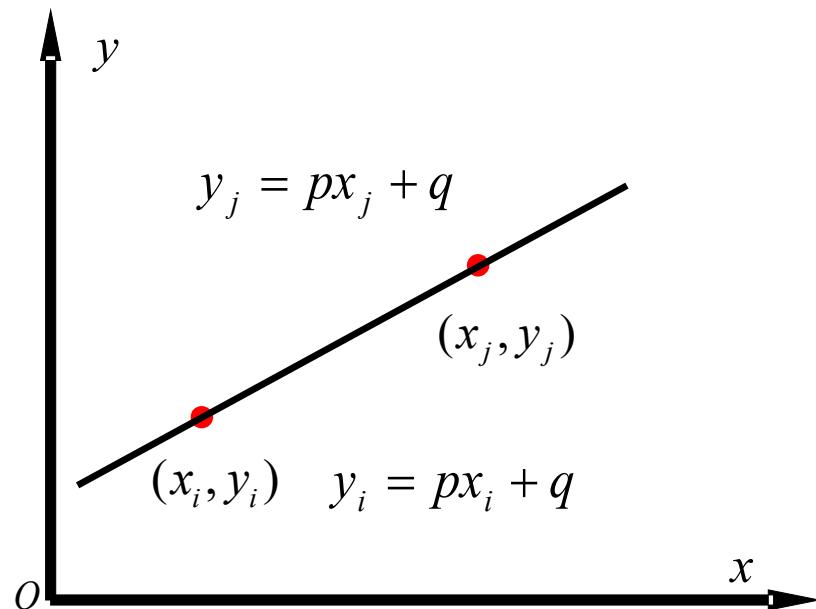
(p,q) 所在空间定义为参数空间 PQ 。



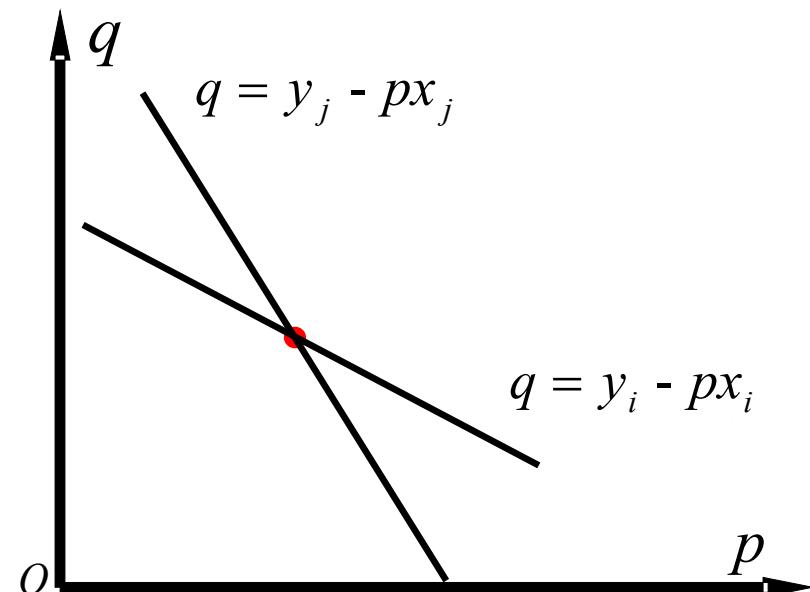
图像 XY 空间

对偶

参数 PQ 空间



(a) □□□□□□□□□



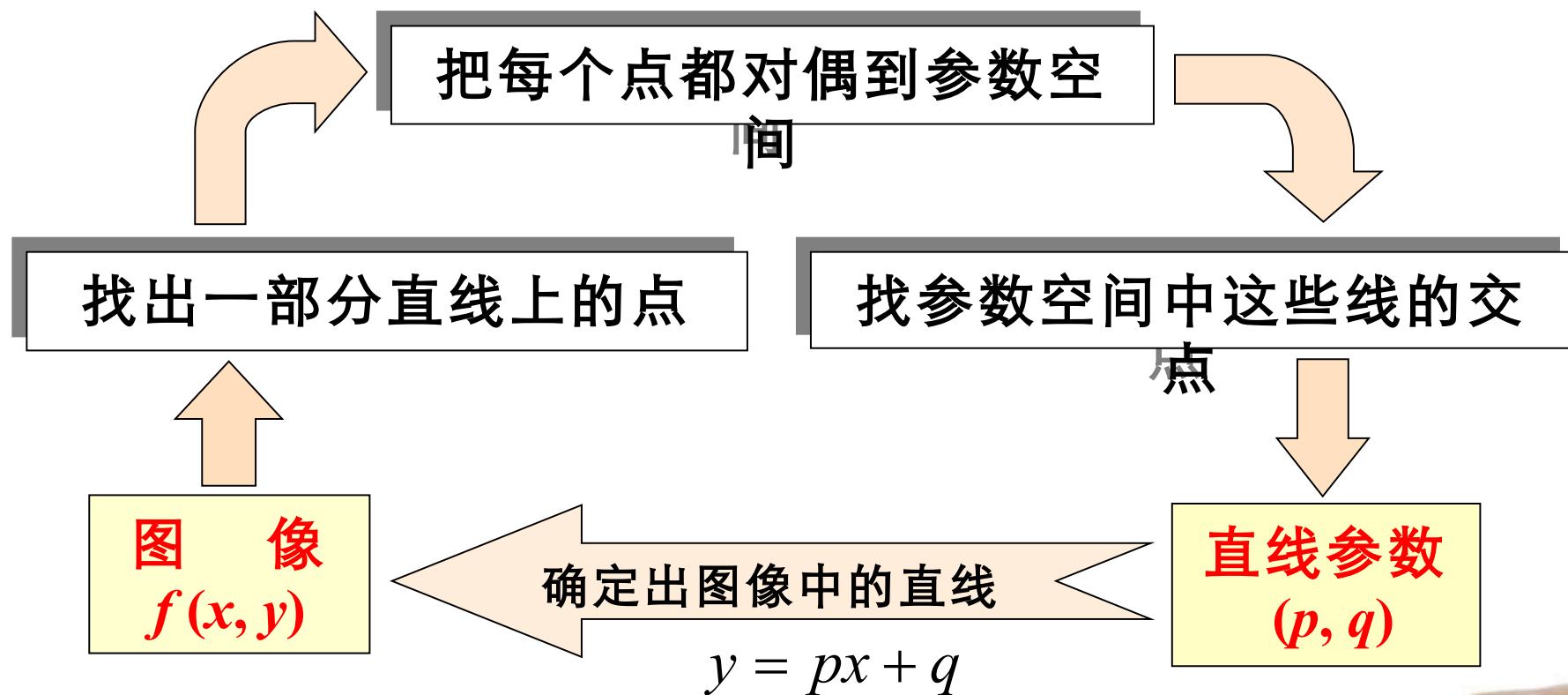
(b) □□□□□□□□□□



$\square \square XY \square \square$		$\square \square PQ \square \square$
$\square \square \square y$		$\square \square q$
$\square \square \square x$		$\square \square p$
$\square (x, y)$		$\square q = y - px$
$\square \square \square \square$ $\square \square \square \square \square \square \square (p_0, q_0)$		$\square \square \square \square$ $\square \square \square \square (p_0, q_0)$



Hough 变换检测直线的基本原理





3、算法实现

准备阶段

估计斜率 p 和截距 q 的范围，并进行离散化
[$p_1, p_2, p_3, \dots, p_m$], [$q_1, q_2, q_3, \dots, q_n$]

建立一个 $m \times n$ 累计数组(p,q) 并清零。

运算阶段

对任一点 (x_i, y_i) ，让 p 遍历 $[p_1, p_2, \dots, p_m]$
对任一 p_j ，按照 $y_i - p_j x_i$ ，求出对应的
累计数组 $A(j, k)$ 加 1。

寻峰阶段

已知图像中有 t 条直线，则在累计数组中找前 t 个最大值 $A(r, s)$ ， $r, s = 1, 2, \dots, t$
则图像中的直线为 $y = p_{u_r} x + q_{v_s}$ 。

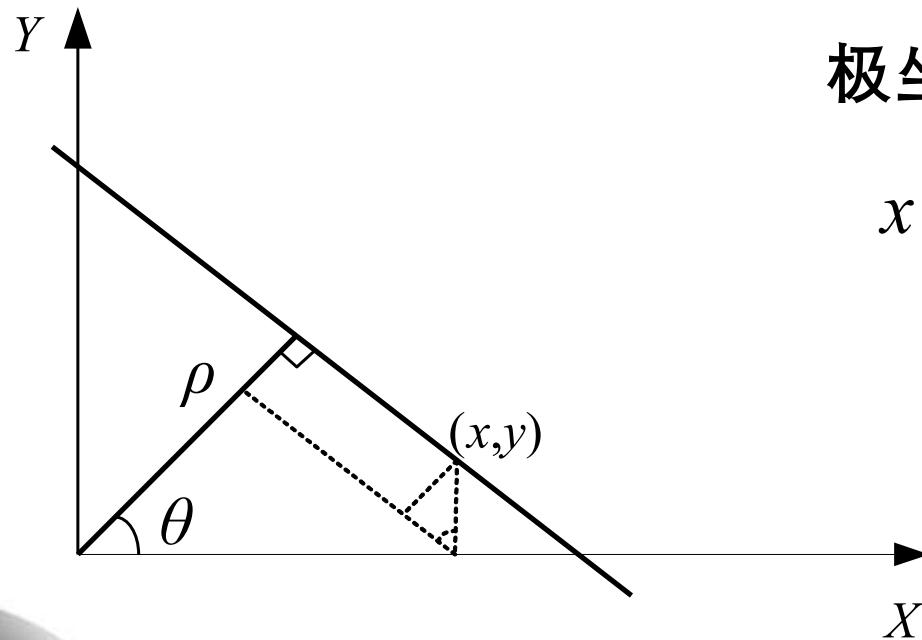


存在的问题及改进

存在的问题：

当直线接近垂直时，斜率 p 接近无穷大，搜索范围很宽，运算量大。

解决的办法 **图像 XY 空间变换到极坐标空间。**

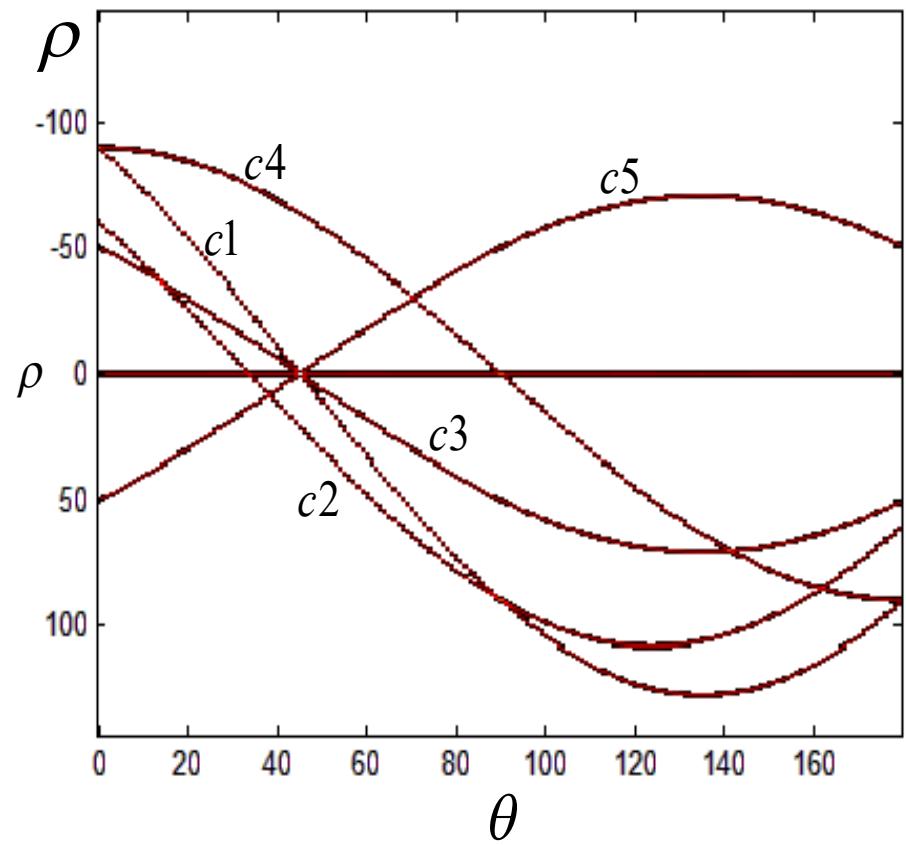
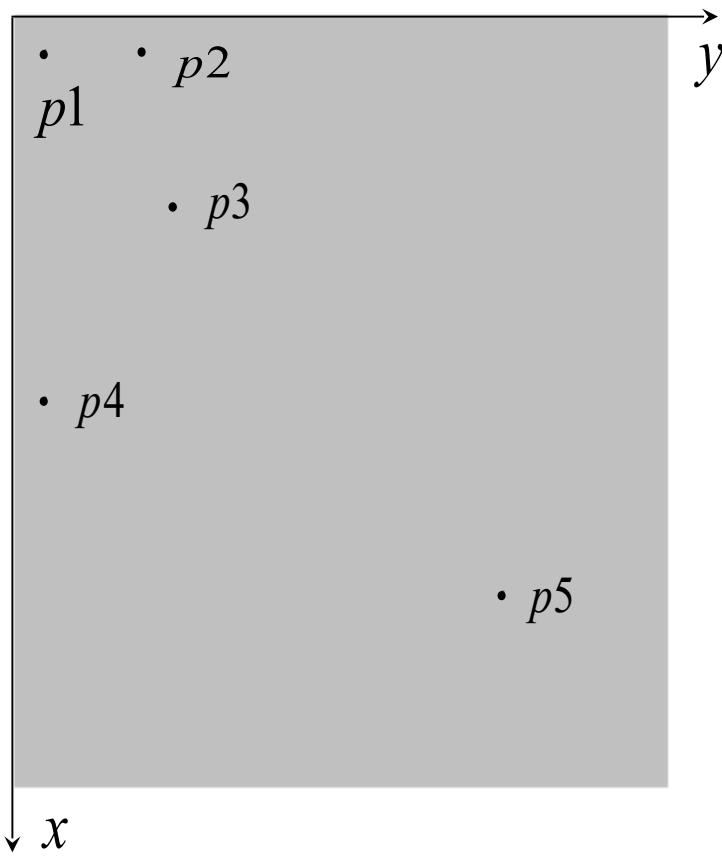


极坐标空间直线方程为：

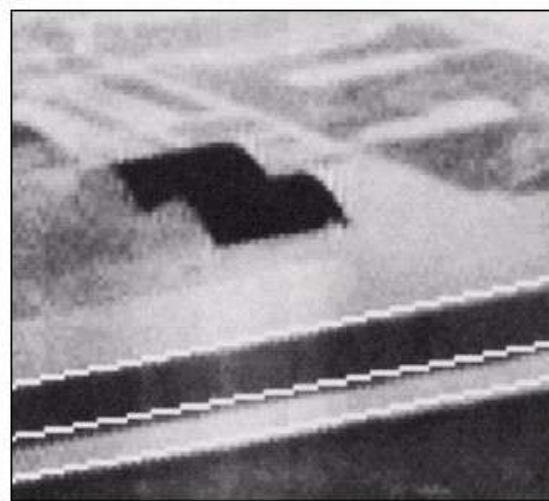
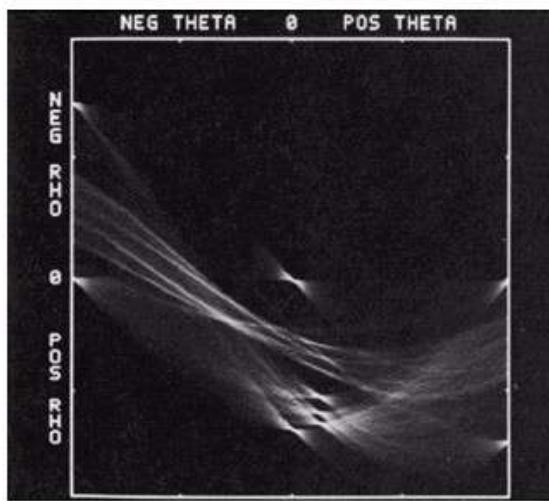
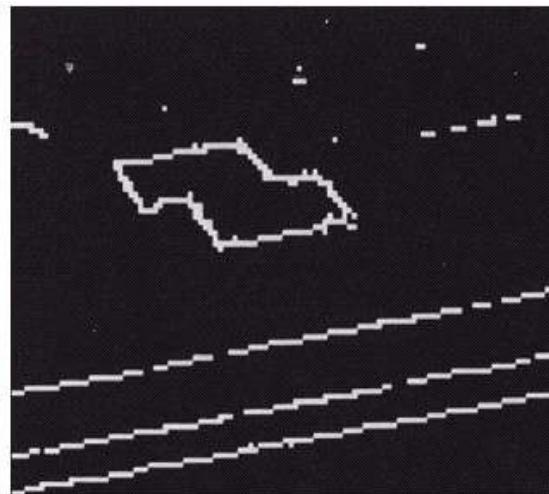
$$x \cos \theta + y \sin \theta = \rho$$



$\square \square XY \square \square$	对偶	$\square \square \rho \theta \square \square$
$\square (x, y)$	对偶	$\square \square \square \rho = x \cos \theta + y \sin \theta$
$\square \square \square \square$ (ρ_0, θ_0)	对偶	$\square \square \square \square \square$ (ρ_0, θ_0)



(a) 航拍的红外线图像 (b) 梯度阈值化图像



(c) 梯度图像的 Hough 变换

(d) Hough 变换检测出的直线



Hough 变换检测直线的特点

- 1) 抗干扰能力强：在参数空间的累加计数是一个积分；
- 2) 鲁棒性好：在真正待求的参数点处有非常明显的尖峰；
- 3) 自动闭合：由边界曲线的参数可得整个曲线。
因此：利用 Hough 变换可排除干扰并完成曲线补充和接续。



4、Hough 变换的扩展

1) 用于圆的检测

圆的方程为： $(x - a)^2 + (y - b)^2 = R^2$

它对应一个 (a, b, R) 的参数空间，计算量大增。

解决途径：加入导数对参数的约束。

对圆的方程求导数，有： $2(x - a) + 2(y - b) \cdot \frac{dy}{dx} = 0$

实际只需二个参数的空间，如 (a, R) 或 (b, R)

。



2) 用于椭圆的检测

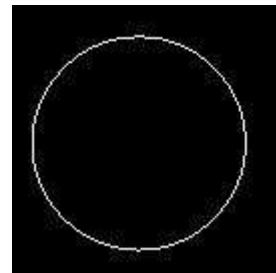
椭圆的方程为： $\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} = 1$

其导数为： $\frac{x - x_0}{a^2} + \frac{y - y_0}{b^2} \cdot \frac{dy}{dx} = 0$

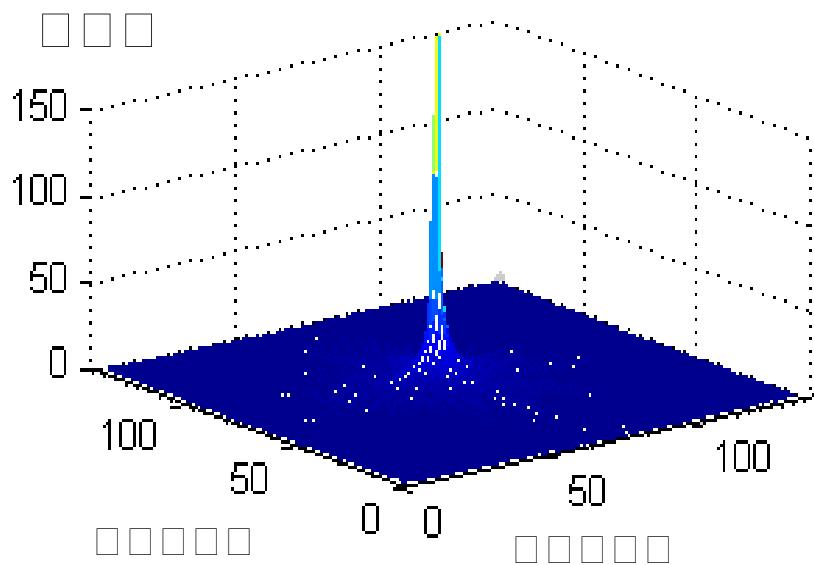
对应的是三参数空间，可随意从 (a, b, x_0, y_0) 中选择三个参数即可。



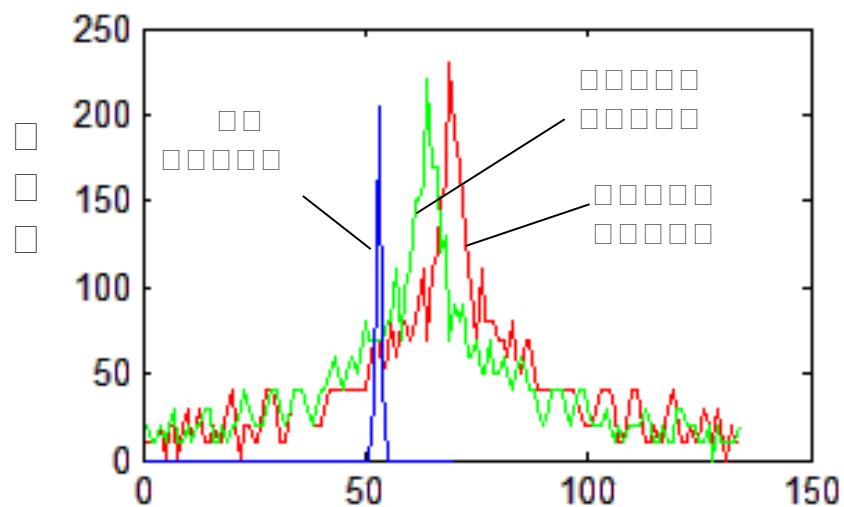
(a) □ □



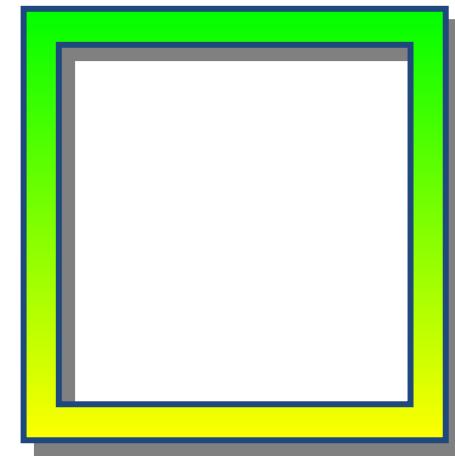
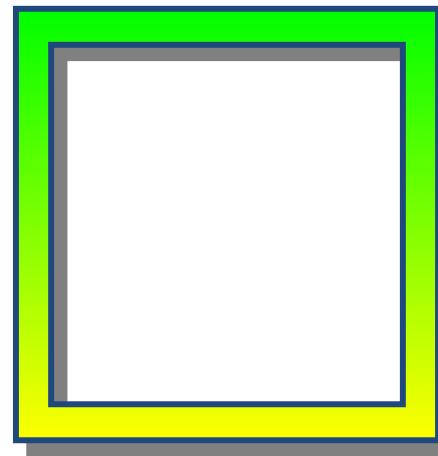
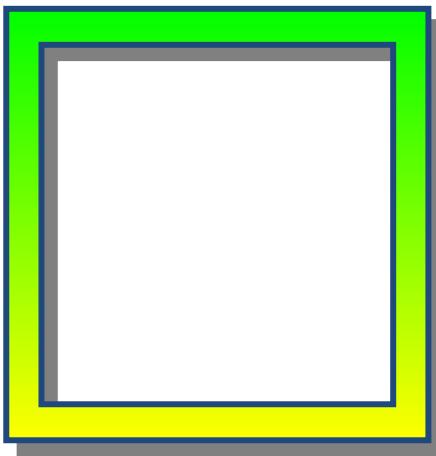
(b) □ (a) □ □ □



(c) □ □ □ □ □ □ □ □ □ □ □



(d) □ □ □ □ □ □ □ □ □ □ □





数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第 7 章 图像描述

第 1 讲 图像描述概述、边界描述（曲线拟合、链码）

演示文稿说明：

- 本讲内容以板书为主，*ppt* 演示为辅；
- 本讲部分图片来自冈萨雷斯的数字图像处理（英文版）教材。



第7章 图像描述

7.1 概述

7.2 边界的表示与描述

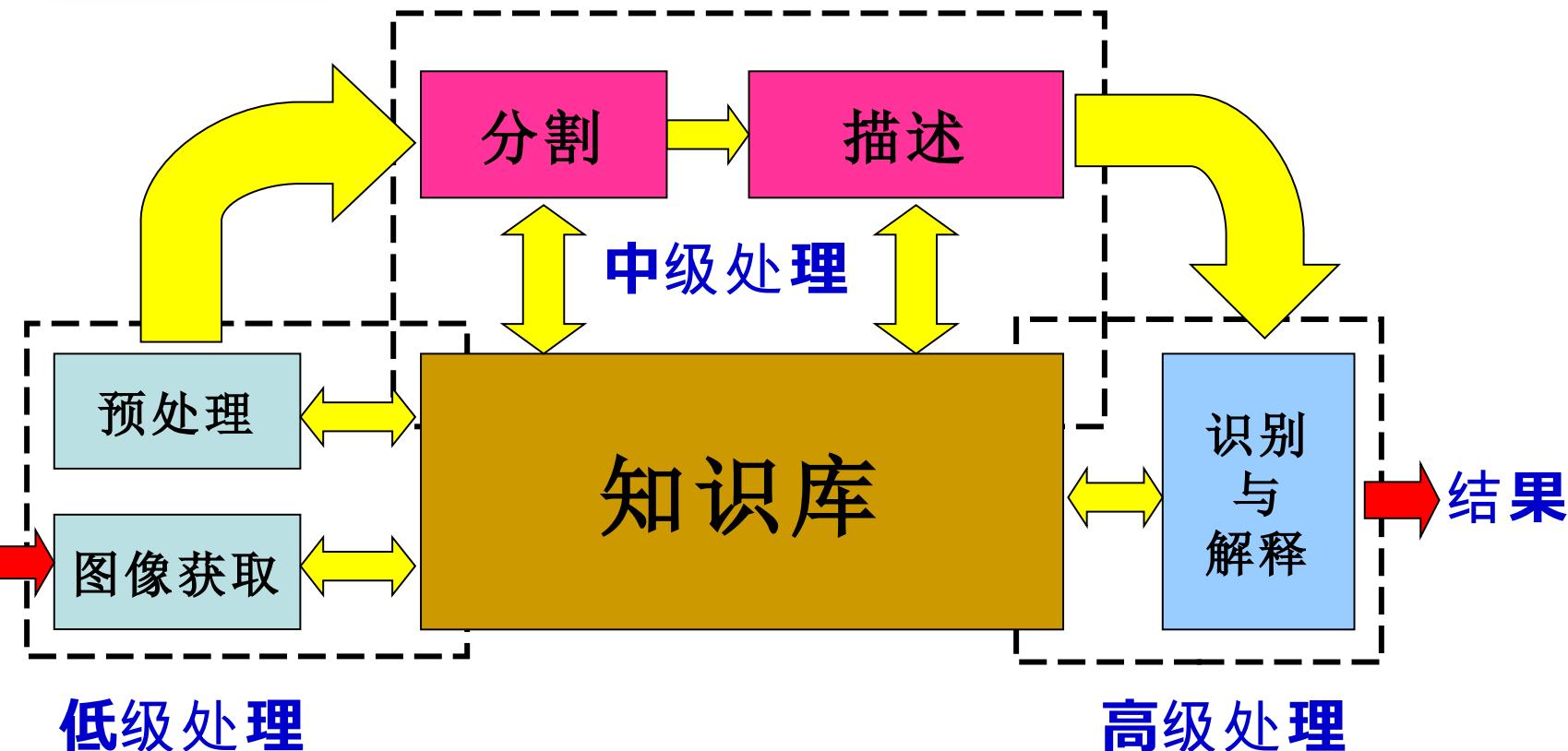
7.3 区域描述



第7章 图像描述

7.1 概述

问题



图像分析系统的基本构成



第7章 图像描述

7.1 概述

- ✓ 图像分割结果是得到了区域内的像素集合，或位于区域边界上的像素集合。
- ✓ 与分割类似，图像中的区域可用其内部（如组成区域的像素集合）表示，也可用其外部（如组成区域边界的像素集合）表示。
- ✓ 一般来说，如果关心的是区域的反射性质，如灰度、颜色、纹理等，常用内部表示法；如果关心的是区域形状，则选用外部表示法。
- ✓ 表示是直接具体地表示目标，好的表示方法应具有节省存储空间、易于特征计算等优点。



第7章 图像描述

7.1 概述

- ✓ 描述是较抽象地表示目标。好的描述应在尽可能区别不同目标的基础上对目标的尺度、平移、旋转等不敏感，这样的描述比较通用
- ✓ 描述可分为对边界的描述和对区域的描述。此外，边界和边界或区域和区域之间的关系也常需要进行描述
- ✓ 表示和描述是密切联系的。表示的方法对描述很重要，因为它限定了描述的精确性；而通过对目标的描述，各种表示方法才有实际意义
- ✓ 表示和描述又有区别，表示侧重于数据结构，而描述侧重于区域特性以及不同区域间的联系和差别





第7章 图像描述

7.1 概述

- ✓ 对目标特征的测量是要利用分割结果进一步从图像中获取有用信息，为达到这个目的需要解决两个关键问题：
 - 选用什么特征来描述目标
 - 如何精确地测量这些特征
- ✓ 常见的目标特征分为灰度（颜色）、纹理和几何形状特征等。其中，灰度和纹理属于内部特征，几何形状属于外部特征
- ✓ 描述方法
 - 边界描述
 - 曲线拟合、链码、周长、曲率以及傅立叶描述子等
 - 区域描述
 - 矩、中轴变换以及拓扑描述子等





第7章 图像描述

7.2 边界的表示与描述

✓ 曲线拟合法

对于任何一个感兴趣的图像目标的边界，认为都是平面中的一条曲线，如果能对该曲线拟合一个函数，则这一函数便可用于描述该目标的边界形状。

设目标边界上的一组点为 $(x_i, y_i), i = 1, 2, \dots, N$

$y = f(x)$ 找出

一个函数 $f(x, y)$

,

使给定点集到 的“距离”之和最小。一般采用均方误差最小准则，通过曲线拟合求出最优化的描述函数

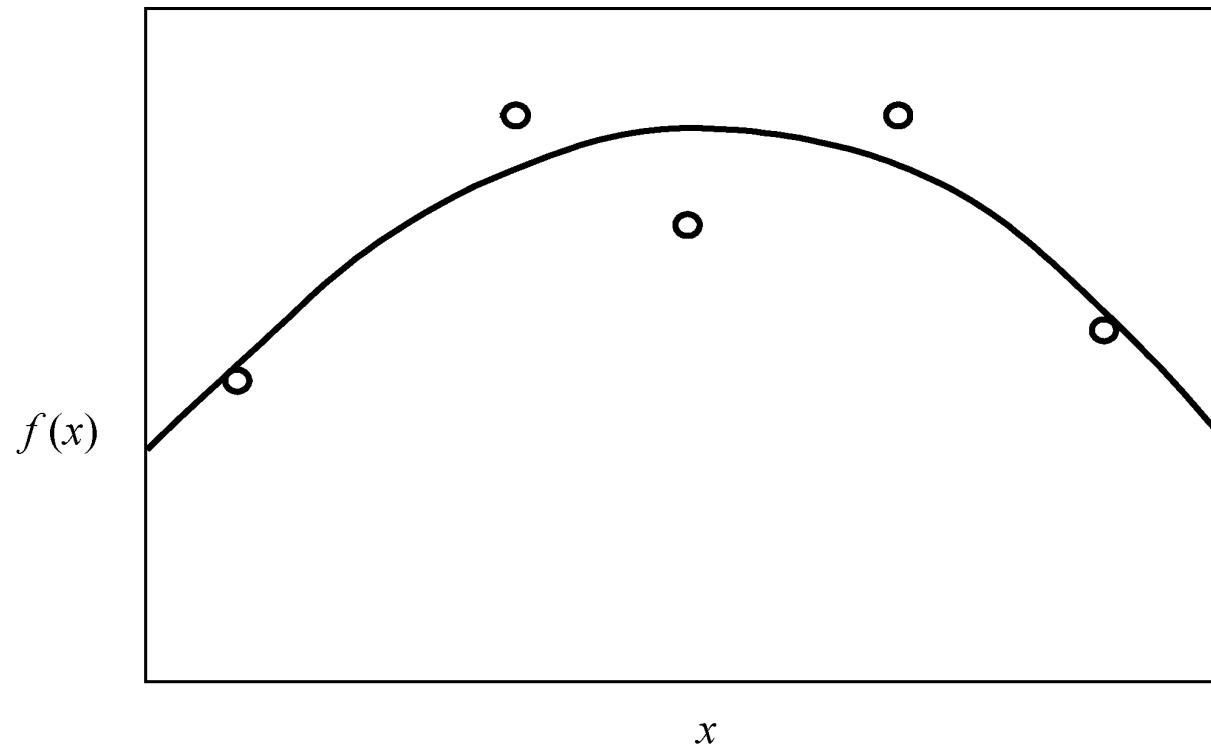
$$MSE = \frac{1}{N} \sum_{i=1}^N [y_i - f(x_i)]^2$$



第7章 图像描述

7.2 边界的表示与描述

✓ 曲线拟合法



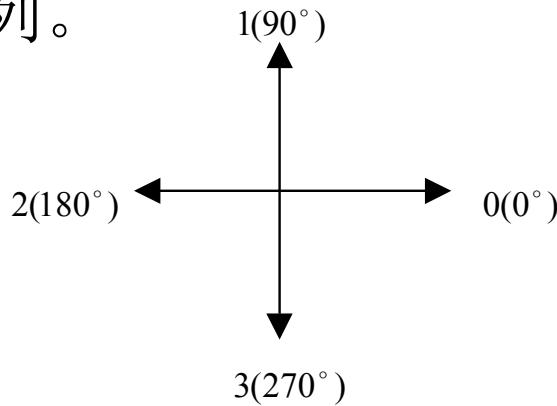


第7章 图像描述

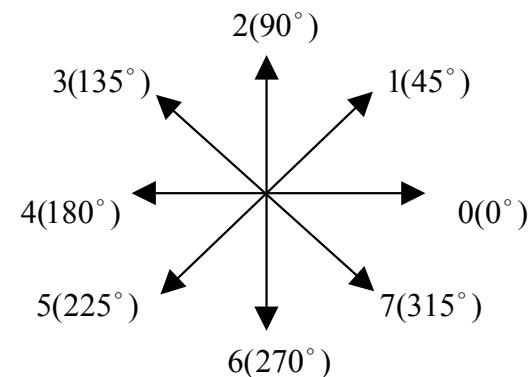
7.2 边界的表示与描述

✓ 链码表示

链码是对边界点的一种编码表示方法，从某一起点开始，跟踪边界并赋给每两个相邻像素的连线一个方向值，结果形成一个数列，可以说，链码实际上是一串方向值的序列。



(a) 4 向链码

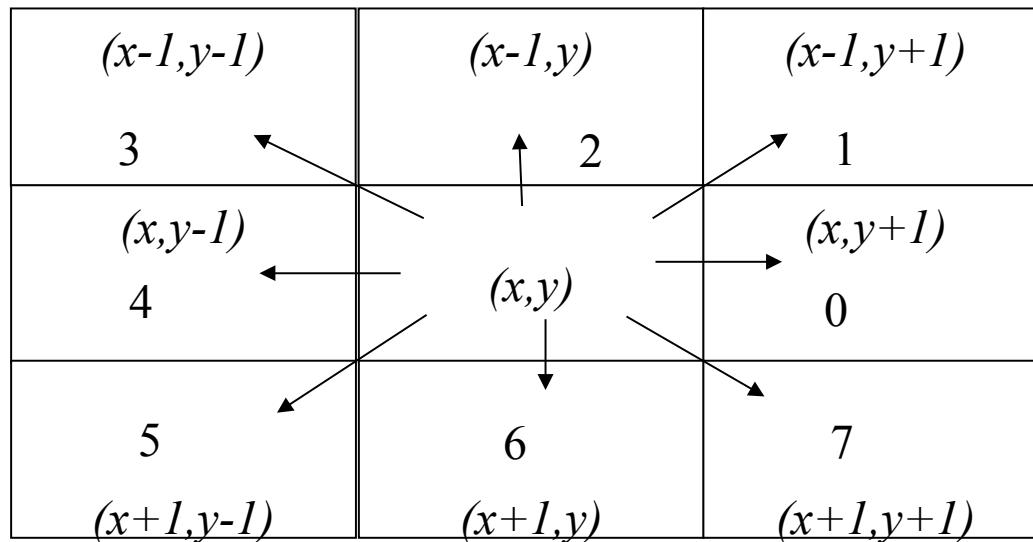
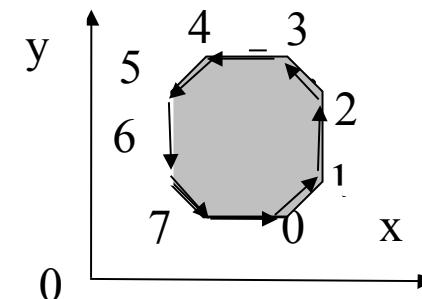
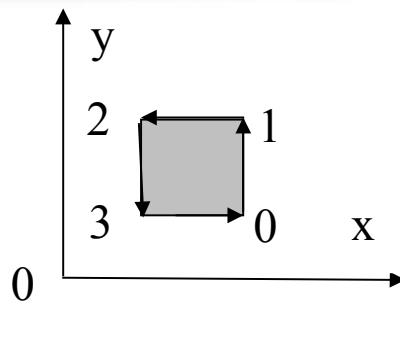


(b) 8 向链码



第7章 图像描述

7.2 边界的表示与描述



链码与坐标位置的关系

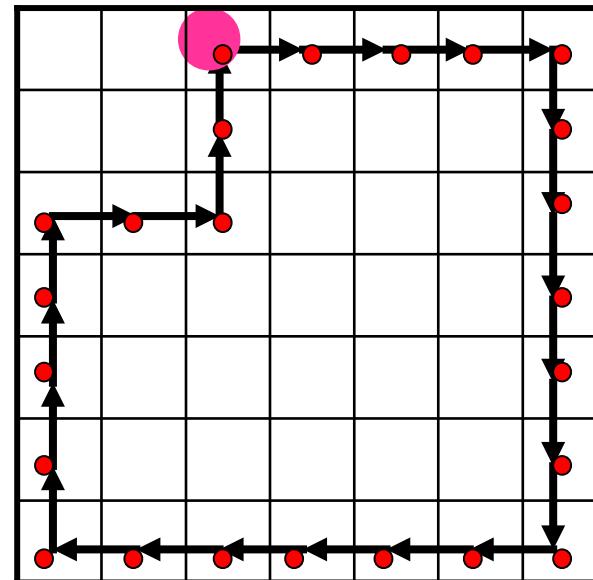




第7章 图像描述

7.2 边界的表示与描述

✓ 链码表示



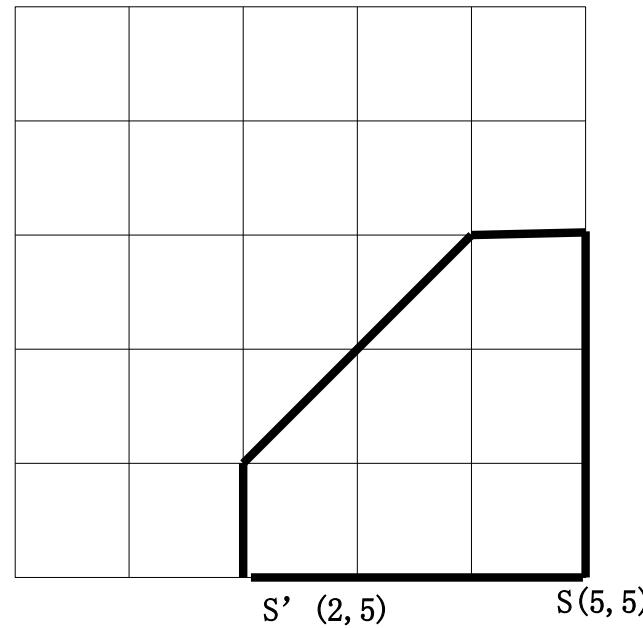
4-链码: 000033333322222211110011



第 7 章 图像描述

7.2 边界的表示与描述

✓ 链码表示



设起始点 s 的坐标为 $(5, 5)$ ，逆时针
分别用 4 方向链码和 8 方向链码表示区域边界：

4 方向链码： $(5, 5) 1 \ 1 \ 1 \ 2 \ 3 \ 2 \ 3 \ 2 \ 3 \ 0 \ 0 \ 0$

8 方向链码： $(5, 5) 2 \ 2 \ 2 \ 4 \ 5 \ 5 \ 6 \ 0 \ 0 \ 0 \ \blacklozenge$



数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第7章 图像描述

第2讲 边界描述(链码、周长、斜率以及傅立叶描述子)

演示文稿说明：

- 本讲内容以板书为主，*ppt* 演示为辅；
- 本讲部分图片来自冈萨雷斯的数字图像处理（英文版）教材。

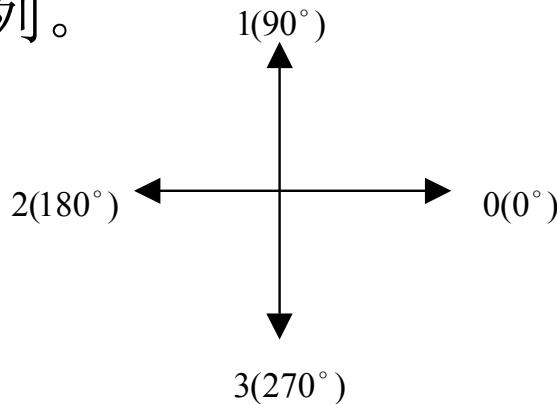


第7章 图像描述

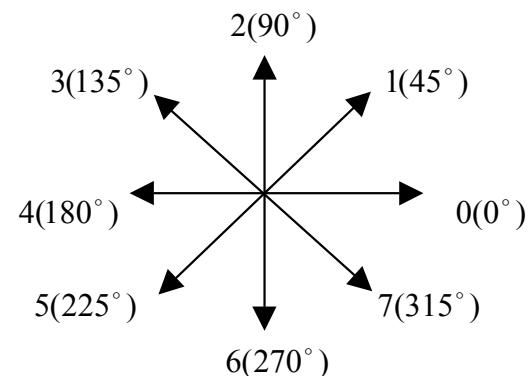
7.2 边界的表示与描述

✓ 链码表示

链码是对边界点的一种编码表示方法，从某一起点开始，跟踪边界并赋给每两个相邻像素的连线一个方向值，结果形成一个数列，可以说，链码实际上是一串方向值的序列。



(a) 4 向链码



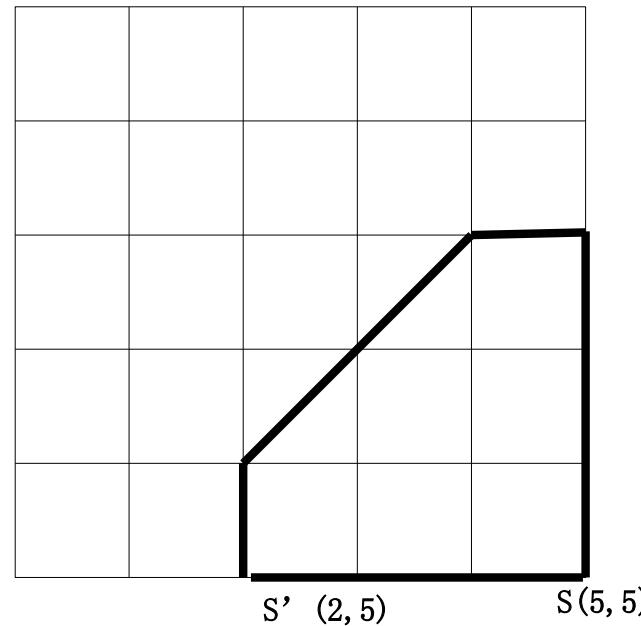
(b) 8 向链码



第 7 章 图像描述

7.2 边界的表示与描述

✓ 链码表示



设起始点 s 的坐标为 $(5, 5)$ ，逆时针
分别用 4 方向链码和 8 方向链码表示区域边界：

4 方向链码： $(5, 5) 1 \ 1 \ 1 \ 2 \ 3 \ 2 \ 3 \ 2 \ 3 \ 0 \ 0 \ 0$

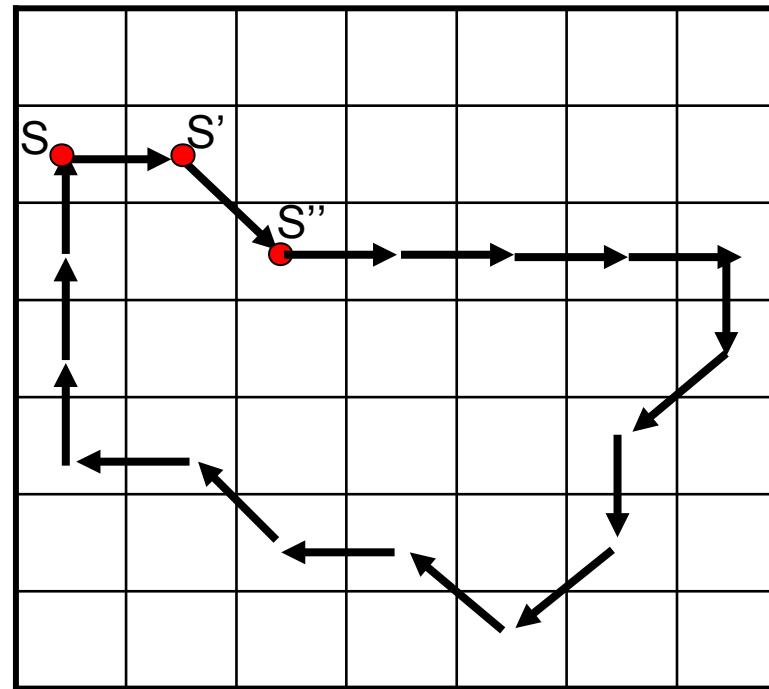
8 方向链码： $(5, 5) 2 \ 2 \ 2 \ 4 \ 5 \ 5 \ 6 \ 0 \ 0 \ 0 \ \blacklozenge$



第7章 图像描述

7.2 边界的表示与描述

✓ 链码表示



链码问题：

由于起点的不同造成编码的不同？一起点归一化
由于角度的不同造成编码的不同？一差分链码



第7章 图像描述

7.2 边界的表示与描述

- ✓ 链码表示（差分链码）

差分链码：

利用链码的一阶差分来重新构造一个表示原链码各段之间方向变化的新序列，这相当于把链码进行旋转归一化。

过程：

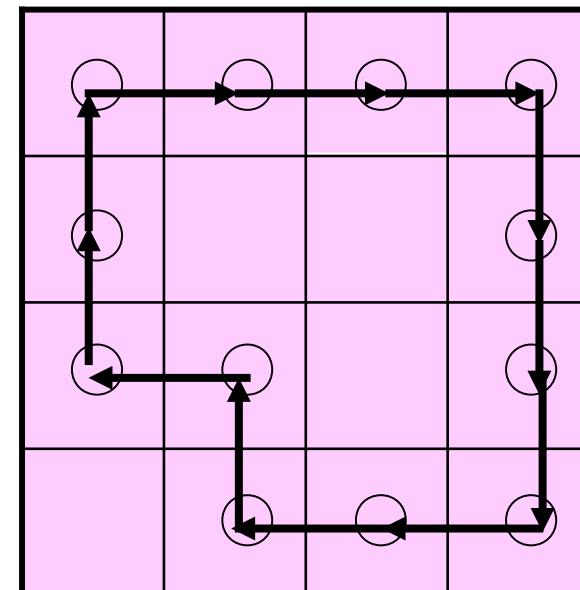
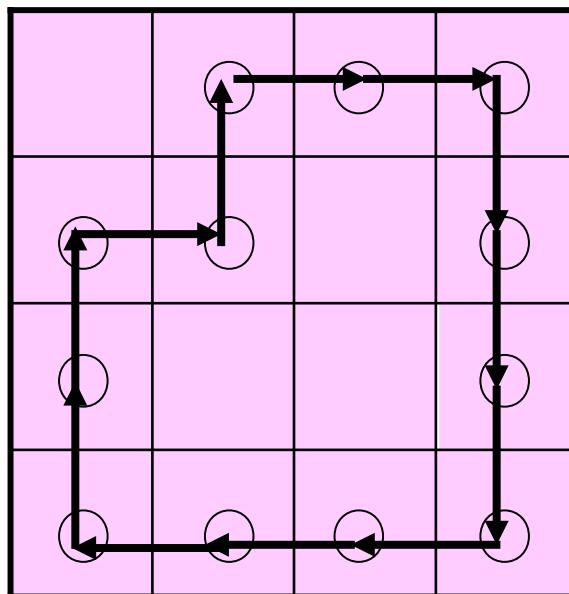
相邻两个方向数按反方向相减（后一个减去前一个），并对结果作模 8（模 4）运算得到。



第7章 图像描述

7.2 边界的表示与描述

- ✓ 链码表示（差分链码）





第 7 章 图像描述

7.2 边界的表示与描述

- ✓ 简单描述子（边界的周长）
 - 区域和背景缝隙的长度和；
 - 8 链码表示：

当链码值为奇数时，其长度记作 $\sqrt{2}$

当链码值为偶数时，其长度记作 1

边界长度 p 表示为 $p = N_e + \sqrt{2}N_o$

式中， N_e 和 N_o 分别是边界链码 (8 方向) 中偶数码元与奇数码元的数目。

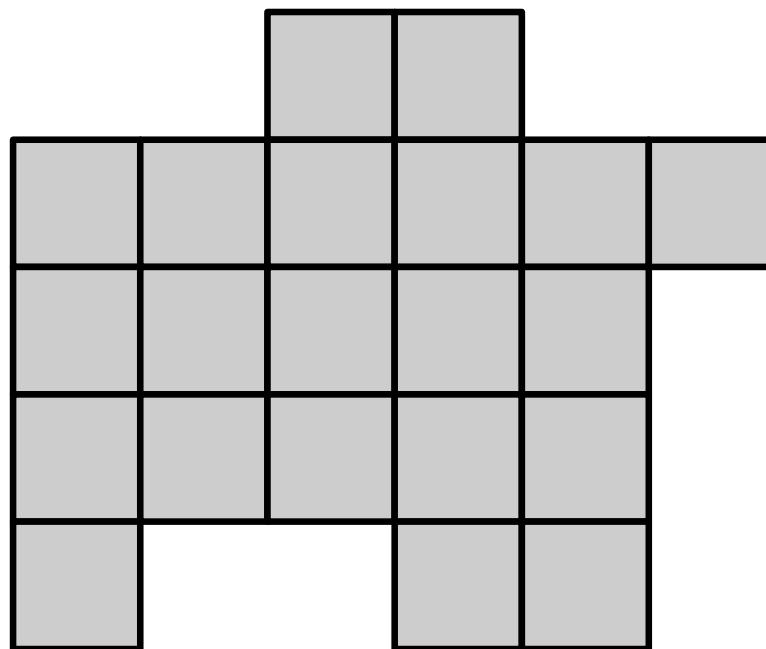
- 边界的像素数



第7章 图像描述

7.2 边界的表示与描述

- ✓ 简单描述子（边界的周长）

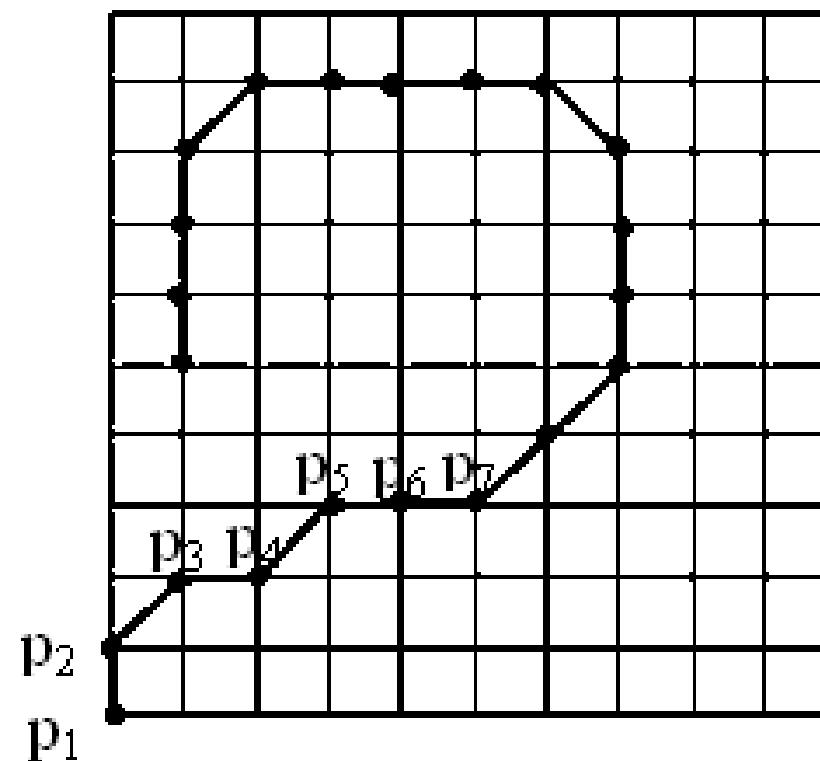




第7章 图像描述

7.2 边界的表示与描述

- ✓ 简单描述子（边界的曲率）





第7章 图像描述

7.2 边界的表示与描述

- ✓ 简单描述子（边界的凸凹线段点）

边界的凸凹点按斜率的符号判定：

斜率非负——凸线段点

斜率为负——凹线段点



第7章 图像描述

7.2 边界的表示与描述

✓ 傅立叶描述子

$$S(k) = x_k + jy_k, k = 0, \dots, N-1$$

$$a(u) = \frac{1}{N} \sum_{k=0}^{N-1} S(k) \exp(-j \frac{2\pi u k}{N}), u = 0, 1, 2, \dots, N-1$$

$$S(k) = \frac{1}{N} \sum_{u=0}^{N-1} a(u) \exp(j \frac{2\pi u k}{N}), k = 0, 1, 2, \dots, N-1$$

$$S'(k) = \frac{1}{N} \sum_{u=0}^{P-1} a(u) \exp(j \frac{2\pi u k}{N}), k = 0, 1, 2, \dots, N-1$$

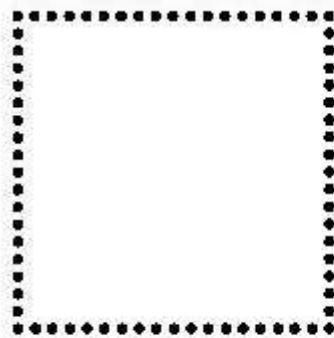


第 7 章 图像描述

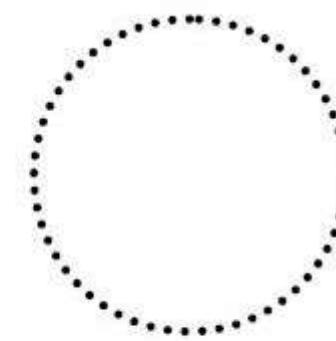
7.2 边界的表示与描述

✓傅立叶描述子 (P 的选取与描述子的关系)

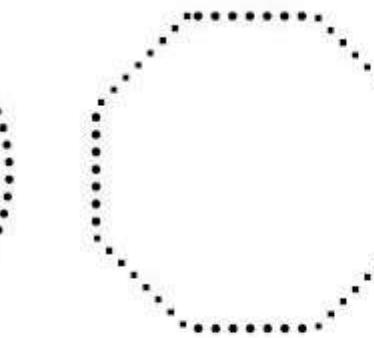
在上述方法中，相当于对于 $u > P-1$ 的部分舍去不予计算。由于傅里叶变换中的高频部分对应于图像的细节描述，因此 P 取得越小，细节部分丢失的越多。



N=64



P=4



P=61



P=62



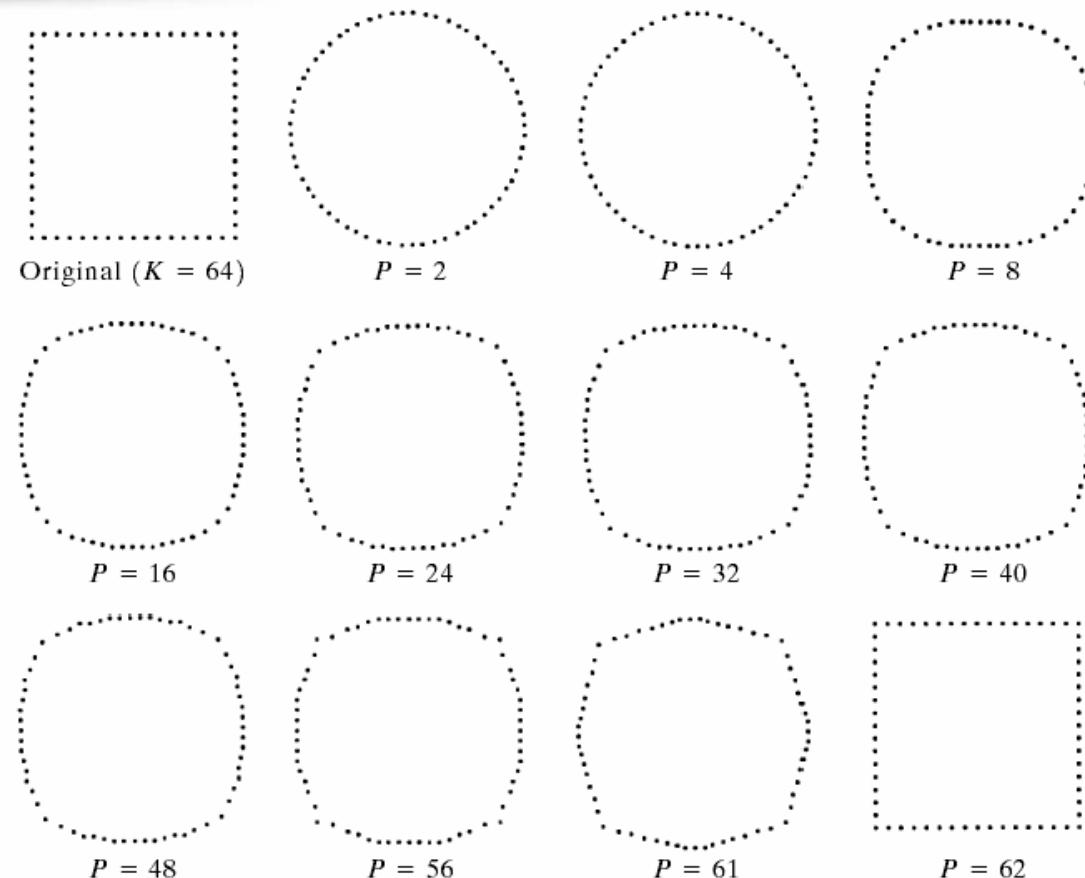


第7章 图像描述

7.2 边界的表示与描述

FIGURE 11.14

Examples of reconstruction from Fourier descriptors. P is the number of Fourier coefficients used in the reconstruction of the boundary.



图片来自冈萨雷斯的数字图像处理(英文版)教材

低阶系数能够反映大体形状，高阶系数可以精确定义形状特征，少数傅里叶描述子携带有形状信息，能够反映边界的大略本质。



第 7 章 图像描述

7.2 边界的表示与描述

- ✓ 傅立叶描述子(使用价值)
 - 较少的傅里叶描述子(如4个)，就可以获取边界本质的整体轮廓。
 - 这些带有边界信息的描述子，可以用来区分明显不同的边界。



数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第7章 图像描述

第3讲 区域描述（骨架、矩描述、简单描述子）

演示文稿说明：

- 本讲内容以板书为主，*ppt* 演示为辅；
- 本讲部分图片来自冈萨雷斯的数字图像处理（英文版）教材。

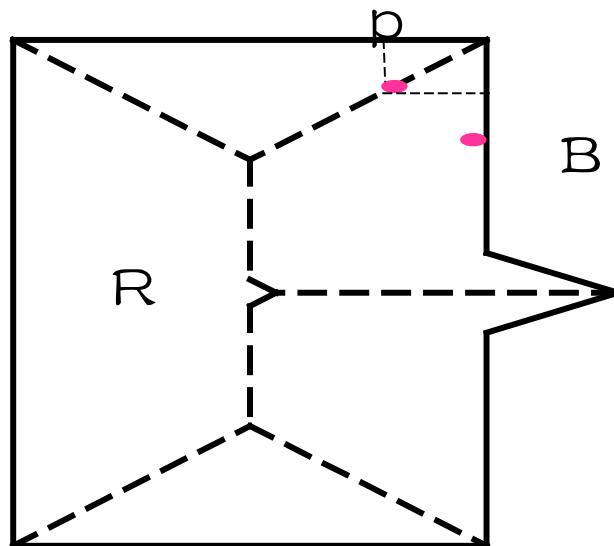


第 7 章 图像描述

7.3 区域描述

✓ 骨架 (Skeleton)

R 是一个区域， B 为 R 的边界点，对于 R 中的点 p ，找 p 在 B 上“最近”的邻居。如果 p 有多于 2 个最近的邻居，称 p 属于 R 的中轴（骨架）

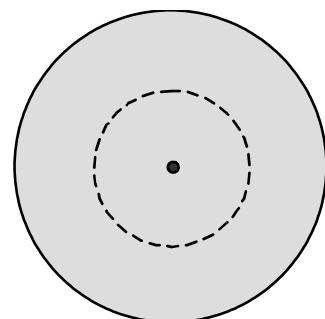




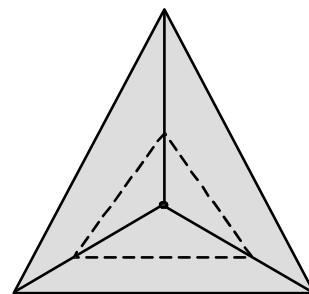
第7章 图像描述

7.3 区域描述

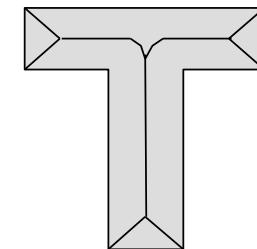
✓ 骨架 (Skeleton)



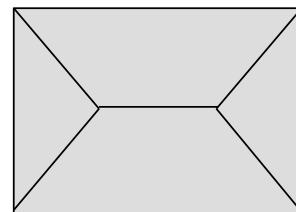
(a)



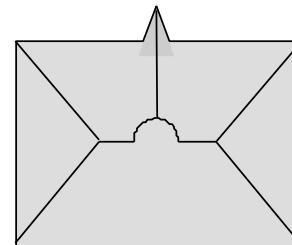
(b)



(c)



(d)



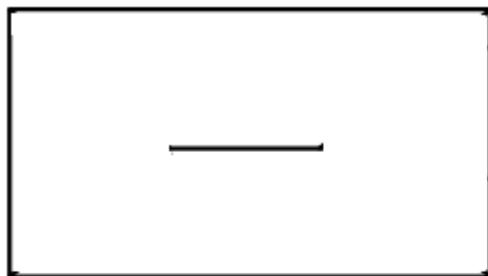
(e)



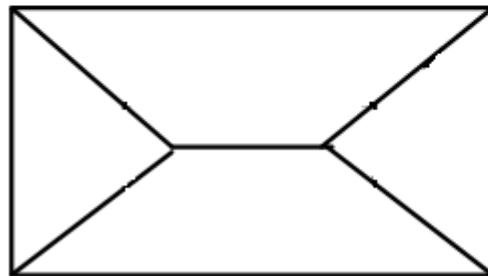
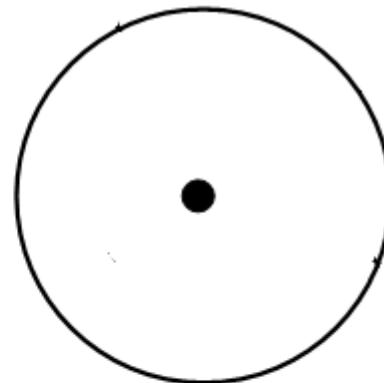
第7章 图像描述

7.3 区域描述

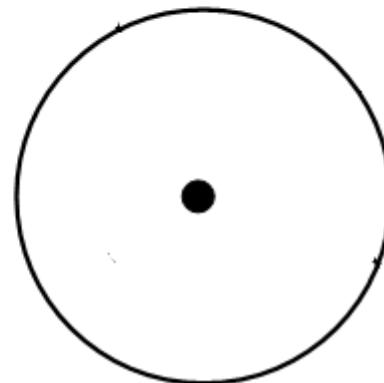
✓ 骨架 (Skeleton)



(a)



(b)





第 7 章 图像描述

7.3 区域描述

- ✓ 矩描述（矩的定义）

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \quad p, q = 0, 1, 2, \dots$$

离散形式：

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad p, q = 0, 1, 2, \dots$$

- ✓ 矩描述（质心坐标与中心矩）

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$$

$$\mu_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy$$

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$



第7章 图像描述

7.3 区域描述

✓ 矩描述

中心矩和原点矩之间存在一定关系 (三阶以内如下)

$$\mu_{00} = m_{00}$$

$$\mu_{02} = m_{02} - \bar{y}m_{01}$$

$$\mu_{10} = 0$$

$$\mu_{30} = m_{30} - 3\bar{x}m_{20} + 2\bar{x}^2m_{10}$$

$$\mu_{01} = 0$$

$$\mu_{03} = m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2m_{01}$$

$$\mu_{11} = m_{11} - \bar{y}m_{10}$$

$$\mu_{21} = m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}^2m_{01}$$

$$\mu_{20} = m_{20} - \bar{x}m_{10}$$

$$\mu_{12} = m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}^2m_{01}$$

归一化的中心矩表示为 η_{pq} :

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad \text{其中 } \gamma = \frac{p+q}{2} + 1 \quad \text{for } p+q = 2, 3, \dots$$



第7章 图像描述

7.3 区域描述

- ✓ 矩描述（不变矩）

$$\phi_1 = \eta_{20} + \eta_{02}$$

$$\phi_2 = (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] +$$

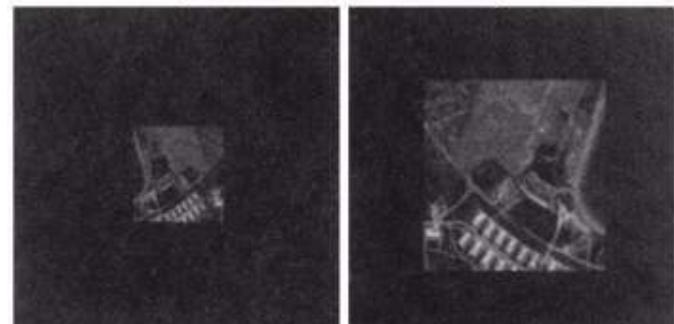
$$(3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$



第 7 章 图像描述

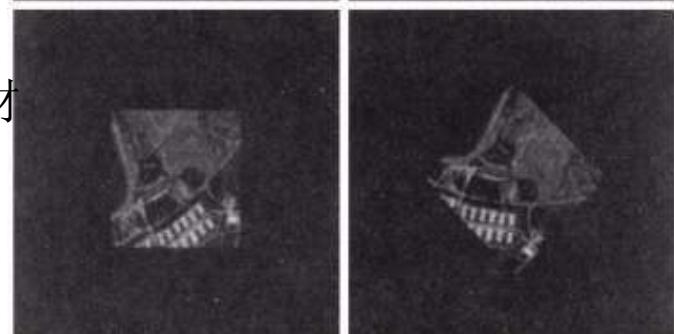
7.3 区域描述

✓ 矩描述



Invariant (Log)	Original	Half Size	Mirrored	Rotated 2°	Rotated 45°
ϕ_1	6.249	6.226	6.919	6.253	6.318
ϕ_2	17.180	16.954	19.955	17.270	16.803
ϕ_3	22.655	23.531	26.689	22.836	19.724
ϕ_4	22.919	24.236	26.901	23.130	20.437
ϕ_5	45.749	48.349	53.724	46.136	40.525
ϕ_6	31.830	32.916	37.134	32.068	29.315
ϕ_7	45.589	48.343	53.590	46.017	40.470

图片来自冈萨雷斯的数字图像处理（英文版）教材





第7章 图像描述

7.3 区域描述

- ✓ 简单描述子
 - 区域面积
 - 矩形度
 - 圆形度



数字图像处理

(Digital Image Processing)

山东科技大学
曹茂永 教授



第7章 图像描述

第4讲 区域描述（简单描述子、拓扑描述子、纹理描述）

演示文稿说明：

- 本讲内容以板书为主，*ppt* 演示为辅；
- 本讲部分图片来自冈萨雷斯的数字图像处理（英文版）教材。



第7章 图像描述

7.3 区域描述

- ✓ 简单描述子
 - 区域面积
 - 矩形度
 - 圆形度
 - 偏心率



第7章 图像描述

7.3 区域描述

- ✓ 拓扑描述子

拓扑性质：



第7章 图像描述

7.3 区域描述

- ✓ 纹理描述
- 统计方法
- 结构方法
- 频谱法



第 7 章 图像描述

7.3 区域描述

- ✓ 纹理描述（统计方法）

- 基于直方图的统计矩

$$\mu_n(z) = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i)$$

其中 z 为任一灰度级， $p(z_i)$ ， $i=0,1,2,\dots,L-1$ ，为对应的灰度直方图 $n = \sum_{i=0}^{L-1} z_i p(z_i)$

$$\mu_0 = 1 \quad \mu_1 = 0 \quad \mu_2 = \sum_{i=0}^{L-1} z_i^2 p(z_i) - m^2 = \sigma^2$$

构造平滑度描述子 $R = 1 - \frac{1}{1 + \sigma^2(z)}$

$\mu_3(z) = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i)$ 对应平均灰度偏离程度



第7章 图像描述

7.3 区域描述

- ✓ 纹理描述（统计方法）
 - 其它基于直方图的纹理量度

一致度:

$$U = \sum_{i=0}^{L-1} p^2(z_i)$$

平均熵:

$$e = -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$$



第7章 图像描述

7.3 区域描述 ✓ 纹理描述

Texture	Average Intensity	Average Contrast	R	Third Moment	Uniformity	Entropy
Smooth	87.02	11.17	0.002	-0.011	0.028	5.367
Coarse	119.93	73.89	0.078	2.074	0.005	7.842
Periodic	98.48	33.50	0.017	0.557	0.014	6.517

TABLE 11.3
Texture measures
for the regions
shown in
Fig. 11.19.

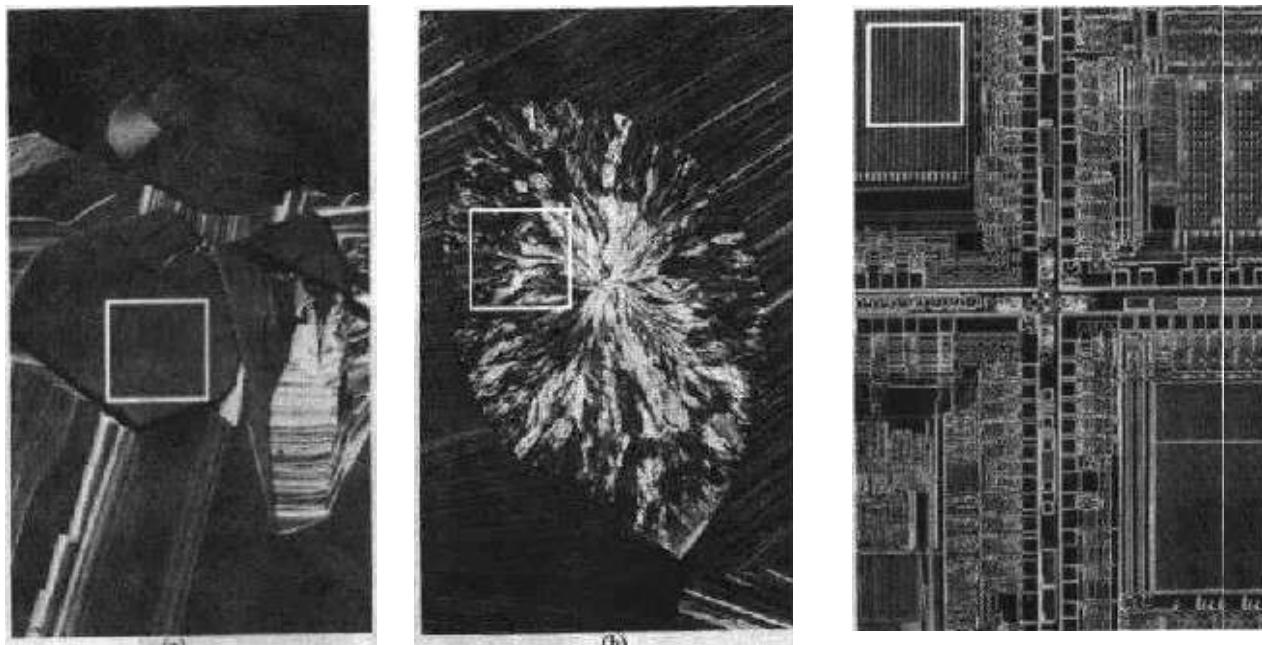


FIGURE 11.19 The subimages shown represent, from left to right, smooth, coarse, and periodic texture. These are optical microscope images of a superconductor, human cholesterol, and a microprocessor. (Original images courtesy of Dr. Michael W. Davidson, Florida State University.)



第7章 图像描述

7.3 区域描述

- ✓ 纹理描述（结构方法）



第7章 图像描述

7.3 区域描述

- ✓ 纹理描述（频谱法）



数字图像处理专题系列

视频监控技术

山东科技大学
赵猛 博士



视频监控技术

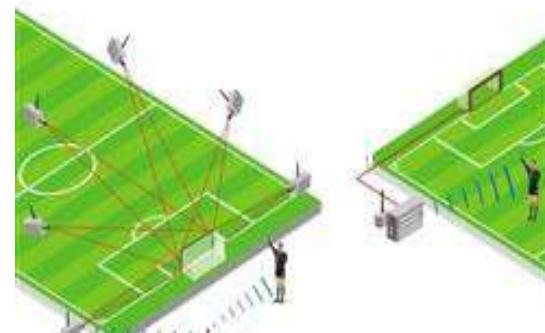
视频监控在军事、银行、交通、安检、企事业以及家庭等领域的安全防范和现场记录报警等方面有着非常广泛的应用。





视频监控技术

视频监控技术是通过计算机对视频信号进行处理、分析和理解，使视频监控系统具有智能。例如在智能交通中违规行驶，地铁、商场场景中客流量的统计，球门线检测是否进球以及在家政服务中老人摔倒等事件的识别。





视频监控技术

视频监控 (Video surveillance) 技术相关内容

运动平台：数字图像稳定 Digital image stabilization

静止平台：

背景维护

Background maintenance

前景分割

Foreground segmentation

底层：视频处理

目标检测

Object detection

目标跟踪

Object tracking

中层：视频分析

动作识别

Action recognition

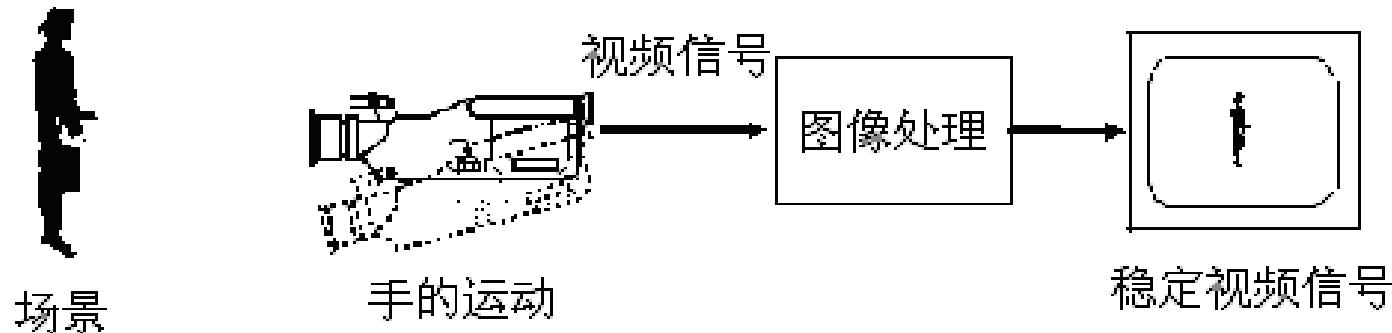
事件识别

Event recognition

高层：视频理解



数字图像稳定 (Digital image stabilization)



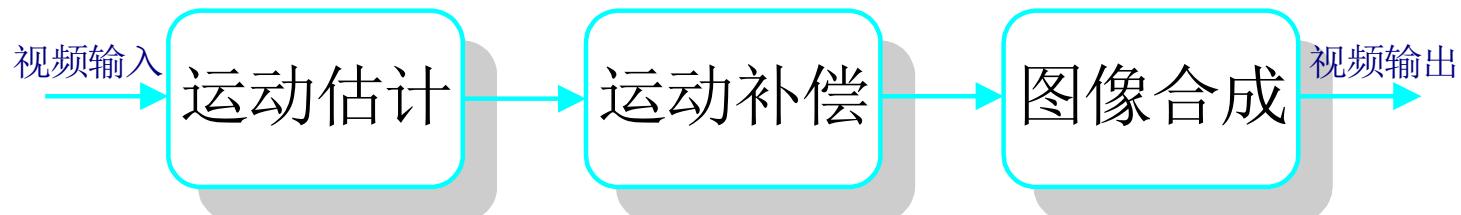


数字图像稳定 (Digital image stabilization)

解决方案：

一般的数字图像稳定系统首先估计摄像机的运动参数，再根据这些运动参数进行运动补偿，最后通过图像的合成得到平滑的视频序列。

- 图像序列的帧间运动估计 (Motion Estimation)
- 运动补偿 (Motion Compensation)
- 图像合成 (Image Composition)





数字图像稳定 (Digital image stabilization)

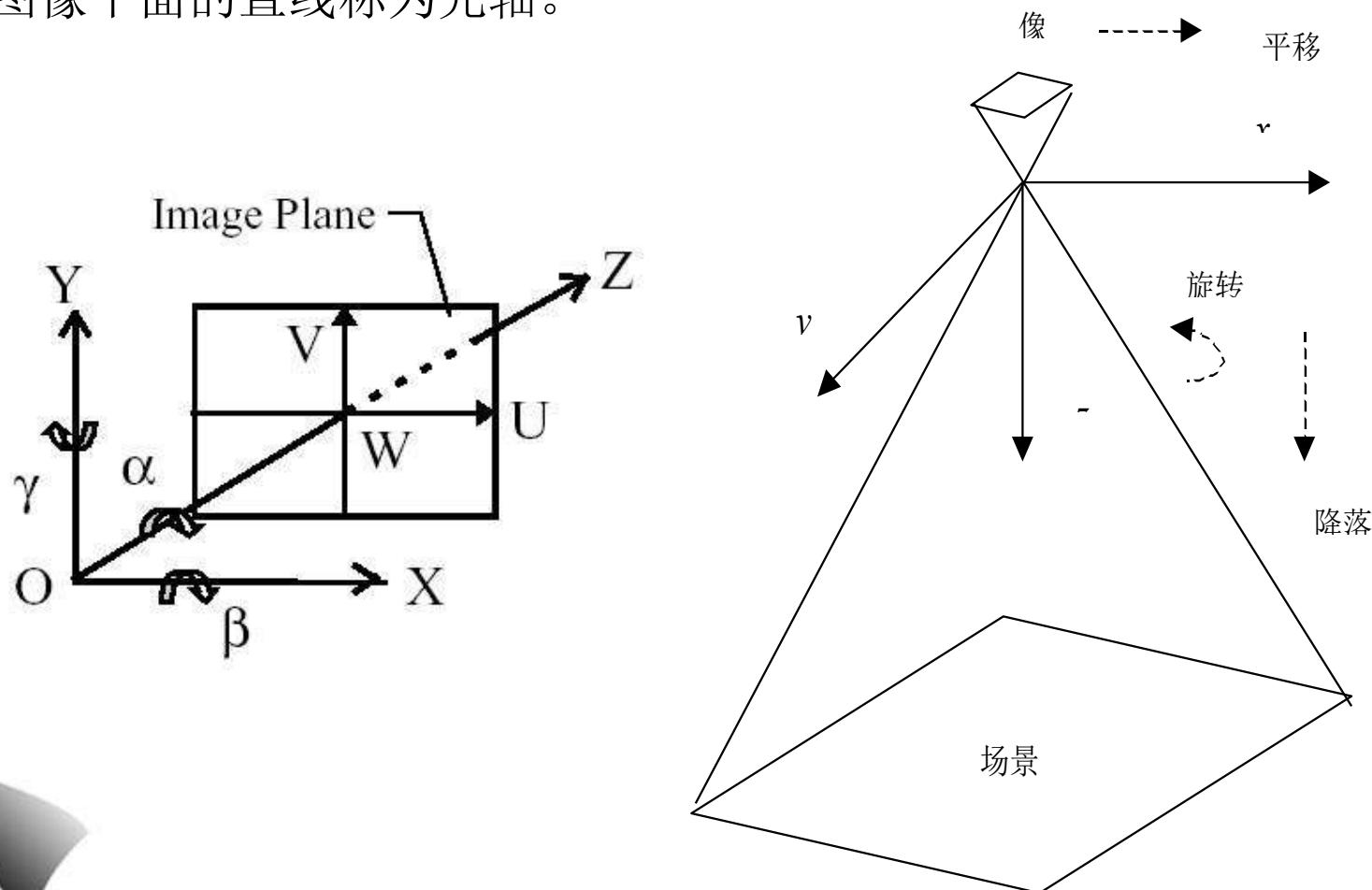
关键技术：

- **运动估计**是估计帧间的运动参数，主要是运动模型的选择和图像运动的检测。首先需要选择一定的运动模型，然后在此基础上计算和估计运动参数。
- **运动补偿**是根据运动估计的结果计算出全局的变换参数，去除不需要的 (unwanted) 运动。
- **图像合成**是根据运动补偿模块的全局变换参数来改变当前帧，得到完整的合成图像。



数字图像稳定 (Digital image stabilization)

透视投影成像模型：透视投影 (Perspective Projection) 可以用针孔 (Pinhole) 成像模型来近似表示。透视投影成像模型的特点是所有来自场景的光线均通过一个投影中心，它对应于透镜的中心，经过投影中心且垂直于图像平面的直线称为光轴。





数字图像稳定 (Digital image stabilization)

成像模型：

假设摄像机的光轴与场景平面垂直，场景是静止的。以运动摄像机为参考建立坐标系，设摄像机的光轴中心为坐标原点，则在摄像机的不同状态下同一点的三维坐标关系为：

$$\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix}$$

$$\begin{pmatrix} X_0 \\ Y_0 \end{pmatrix} = S \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} + \begin{pmatrix} \Delta X \\ \Delta Y \end{pmatrix}$$

$$S = \frac{z_1}{z_0}$$



数字图像稳定 (Digital image stabilization)

运动参数估计：

运动估计的目标就是估计出四个参数：比例变换 (Scale)、旋转角度 (Rotation)、平移量 (Translation)。估计算法应满足实时性的要求。

二维运动估计的方法一般分为三类：

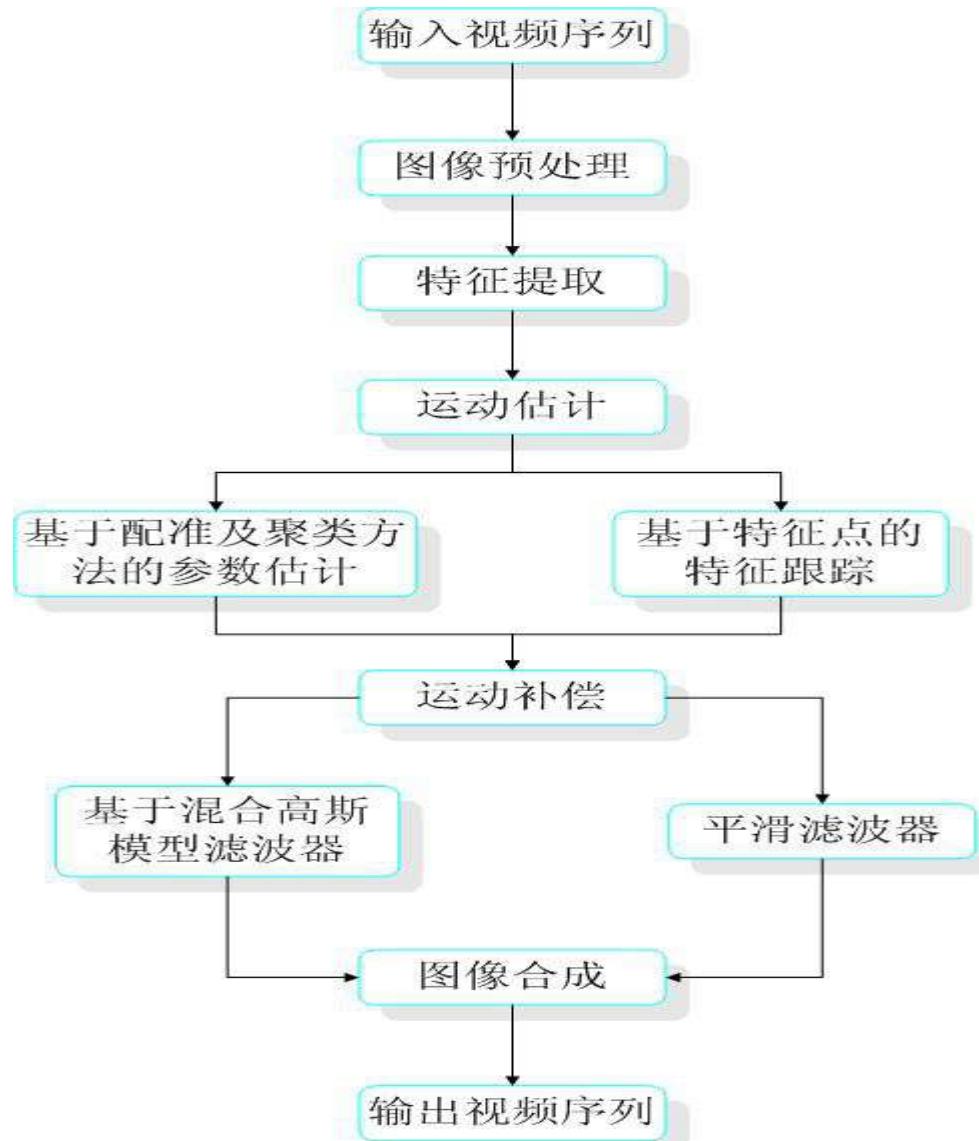
基于光流场：光流场法计算复杂，费时。

基于图像特征：首先对图像特征提取，然后检测图像的特征运动变化，
由图像的特征变化代替整幅图像的运动。

块匹配：从参考图像中找到最好的匹配块来构建当前图像，块匹配方法
分为基于图像和基于图象中的目标的块匹配方法。



数字图像稳定 (Digital image stabilization)





数字图像稳定 (Digital image stabilization)

特征提取：

目标的特征提取的内容包括特征点（如角点）、线、边界、区域等。这里我们介绍[角点](#)的提取方法。

对于一个景物目标，在它的较平直的边缘对应的图像边界或内廓线上梯度方向大致相同，而目标边缘曲率较大的地方或两条、多条边缘的交点所对应的图像中的小区域上的某些相邻点的梯度方向变化较大。两条边界以一定的角度相交的地方、边界方向发生据变的地方以及图像灰度梯度变化较大的地方，我们称这些地方的点为角点。显然，以目标图像的角点作为目标的几何特征点是有重要意义的，因它具有目标几何形状信息。目标的角点可以用于许多目的，如运动估计、形状描述等。



数字图像稳定 (Digital image stabilization)

域相关性角点检测：两幅图像中用于匹配的点应尽可能容易地被识别和匹配。显而易见，一个均匀区域中的点是不适合作为候选匹配点，所以兴趣算子应在图像中寻找具有很大变化的区域。在以某一点为中心的窗函数中，计算其在不同方向上的变化量是这些方向上点的差异性的最好测度。方向变化量的计算公式如下：

$$I_1 = \sum_{(x,y) \in S} [f(x, y) - f(x, y + 1)]^2$$

$$I_2 = \sum_{(x,y) \in S} [f(x, y) - f(x + 1, y)]^2$$

$$I_3 = \sum_{(x,y) \in S} [f(x, y) - f(x + 1, y + 1)]^2$$

$$I_4 = \sum_{(x,y) \in S} [f(x, y) - f(x + 1, y - 1)]^2$$

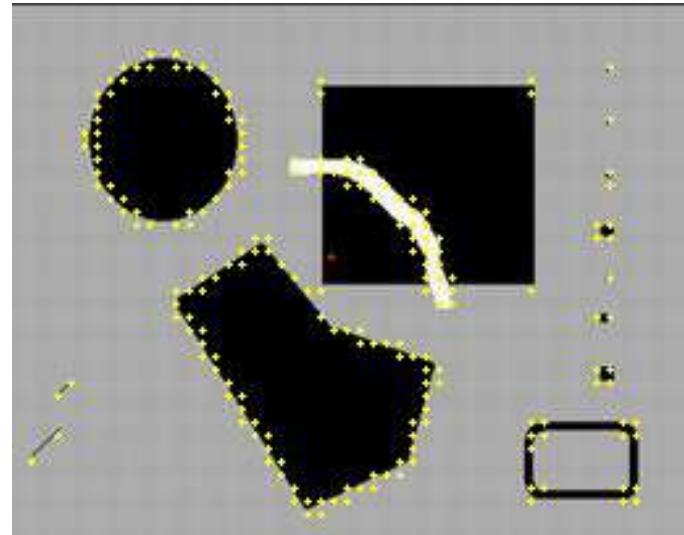
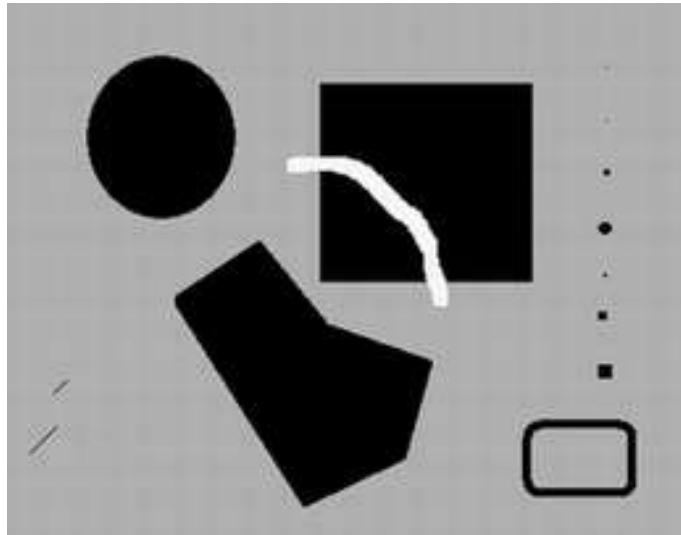
其中 S 表示窗函数中的所有像素。典型窗函数尺寸从 $5*5$ 到 $11*11$ 个像素的范围。因为简单的边缘点在边缘方向上无变化，所以，选择上述方向变量的最小值为中心像素点的兴趣值，可以消除边缘点。

$$I(x_c, y_c) = \min(I_1, I_2, I_3, I_4)$$



数字图像稳定 (Digital image stabilization)

特征点检测:





数字图像稳定 (Digital image stabilization)



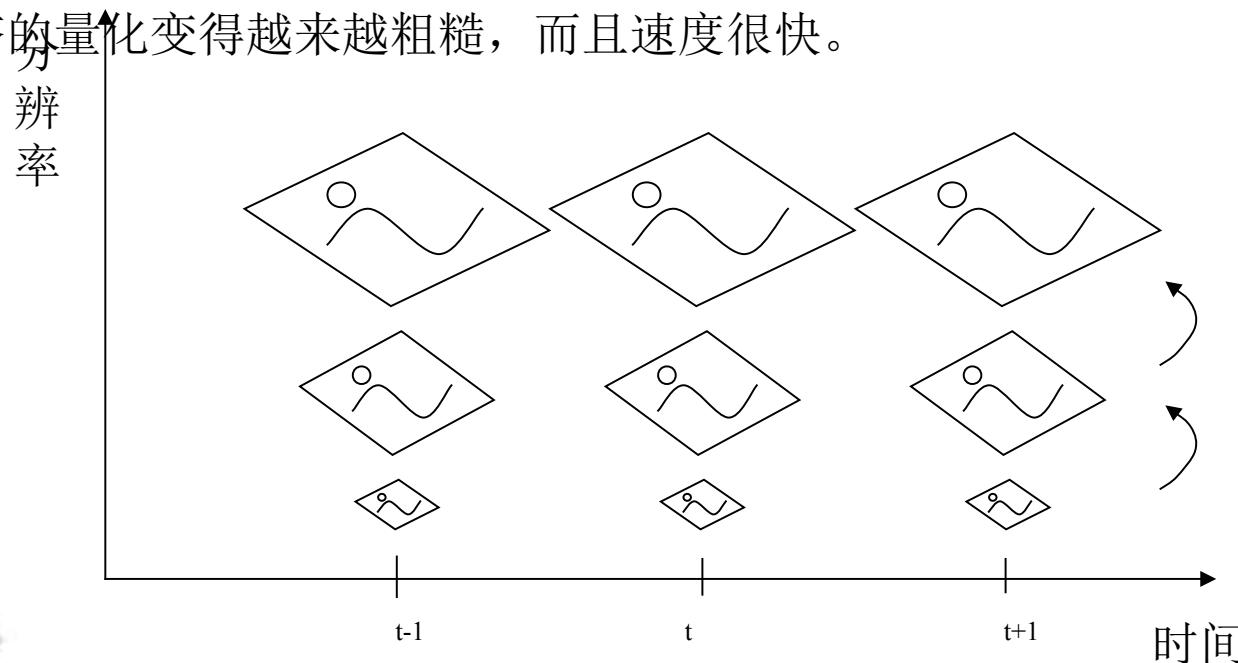


数字图像稳定 (Digital image stabilization)

金字塔技术：图像金字塔是以多分辨率来解释图像的一种结构。（空间换时间）

一幅图像的金字塔是一系列以金字塔形状排列的分辨率逐步降低的图像集合。金字塔的底部是待处理图像的高分辨率表示，而顶部是低分辨率的近似。当向金字塔的上层移动时，尺寸和分辨率就降低。

金字塔是结合降采样操作和平滑操作的一种图像表示方式。它的一个很大的好处是，自下而上每一层的像素数都不断减少，这会大大减少计算量；而缺点是自下而上金字塔的量化变得越来越粗糙，而且速度很快。





数字图像稳定 (Digital image stabilization)

运动补偿 (FFA、FRA、FMA)

帧到帧算法 (Frame-to-Frame Algorithm)

- 以相邻帧间的运动为准，累加的结果用来补偿帧间的运动；
- 没有考虑累计误差的问题。由于后续帧的运动估计可能存在误差，会影响稳定的效果。

帧到参考帧算法 (Frame-to-Reference Algorithm)

- 采用一个参考帧，直接用来估计每一帧的全局运动；
- 参考帧随着时间逐渐孤立，需要限制参数，并自动重置。

帧到合成帧算法 (Frame-to-Mosaic Algorithm)

- 与 FRA 相比较，FMA 是利用多帧信息合成一个参考帧；
- 合成帧随时间变化。



数字图像稳定 (Digital image stabilization)

Demo:

稳定前

original sequence

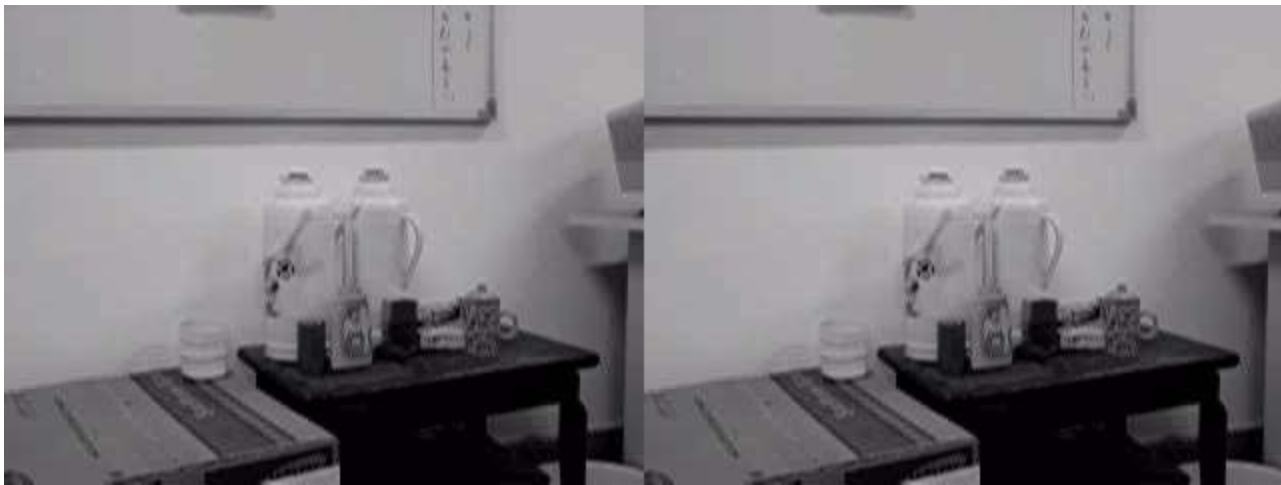
稳定后

stabilized sequence





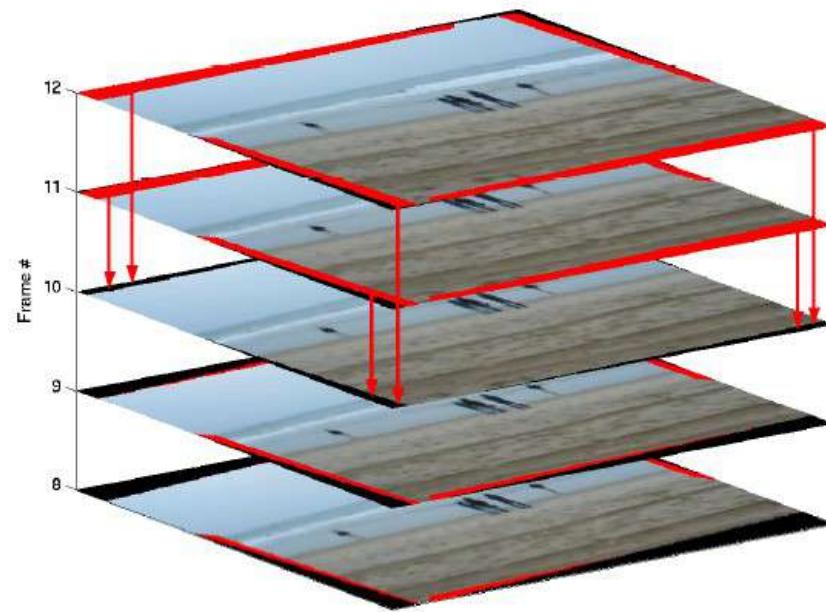
数字图像稳定 (Digital image stabilization)





数字图像稳定 (Digital image stabilization)

图像合成:





数字图像稳定 (Digital image stabilization)

Demo





视频监控技术

静止平台：

背景维护

Background maintenance

前景分割

Foreground segmentation

底层：视频处理

目标检测

Object detection

目标跟踪

Object tracking

中层：视频分析

动作识别

Action recognition

事件识别

Event recognition

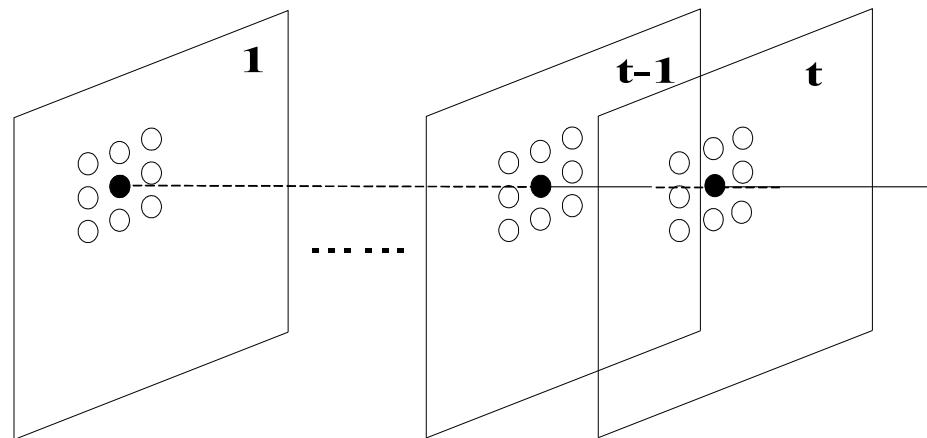
高层：视频理解



背景维护 (BM) 与前景分割 (FS)

背景维护：建立背景模型来生成背景图像的像素值，然后将当前帧与背景图像求差，差值较大的像素区域被认为是运动目标区域，而差值较小的像素区域被认为是背景区域。

$$|B(x) - I_t(x)| > T(x)$$



前景分割：提取出运动目标区域（前景 **Foreground**），能否得到完整的前景区域将对下一步的目标检测有直接的影响。

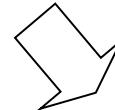


背景维护 (BM) 与前景分割 (FS)

常用的背景模型：混合高斯模型 (GMM)



背景图
像



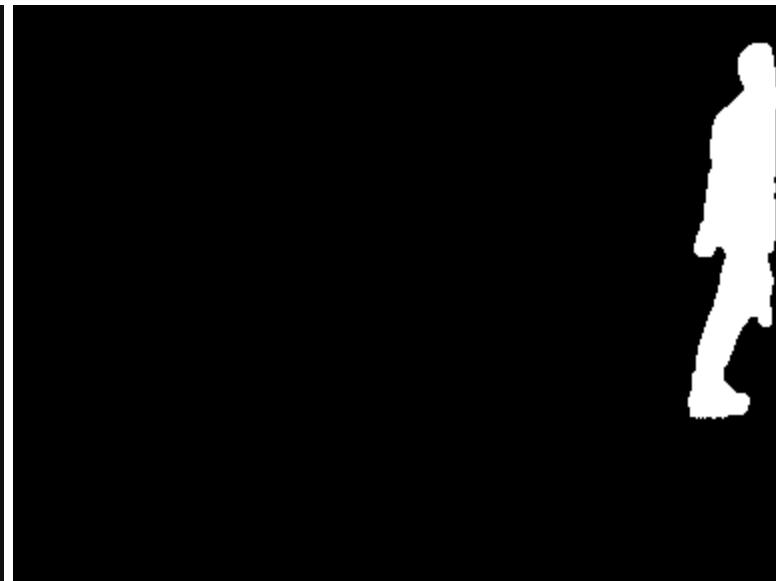
差分后二值化

第 t 帧





背景维护 (BM) 与前景分割 (FS)





背景维护 (BM) 与前景分割 (FS)

背景模型存在的问题：

亮度的变化

lighting change

移动的影子

moving cast shadow

移动场景中的物体

moving background elements of the scene

物体介入或移除

objects introduced or removed from the

scene





目标检测 (Object detection)

目标检测：建立感兴趣的目标模型，从分割出的前景中检测出目标，
用于跟踪。

举 例：人体头部的检测

- Head detection target
 - High detection rate
 - Low false alarm rate
 - Robust to clutter background
 - Working under occlusions
- Head representation
 - Ellipse or circle
 - Certain scale range
- Detection strategy
 - Gradient space
 - Improved Hough transform
 - Hypotheses and verification



目标跟踪 (Object tracking)

视频目标跟踪目的就是对图像序列中的目标进行分析，计算出运动目标的运动参数，如位置、速度、加速度以及运动轨迹等，从而进行进一步处理和分析，实现对运动目标的行为理解，以完成更高一级的视觉任务。

目标跟踪难点：

1. 被跟踪运动目标的姿态变化
2. 运动目标所处环境的光照变化
3. 遮挡问题
4. 实时性要求



目标跟踪 (Object tracking)

目标跟踪的方法：

基于区域的跟踪

利用图像相关匹配技术对目标进行跟踪，需要先获得一个表示目标的模板，模板可以是通过手工选定或是图像分割得到，通常是略大于目标的矩形区域或是任意形状；然后运用相关算法，计算目标模板与当前图像候选区域的特征匹配度，匹配特征可以是颜色、纹理等，匹配度最大的候选区域即为目标的当前位置。

基于主动轮廓的跟踪

基于主动轮廓的方法将目标描述为可变形的边界轮廓来进行跟踪。



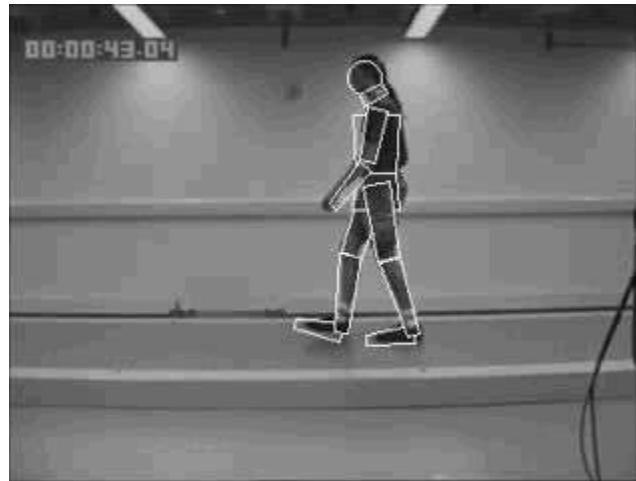
目标跟踪 (Object tracking)

基于特征的跟踪

通过跟踪目标的某个或某些局部特征，如边缘特征、点特征等，从而实现对整个目标的跟踪。

基于模型的跟踪

利用先验知识建立跟踪目标的目标模型（2D、3D 模型），然后将目标模型的图像投影与图像序列进行匹配，从而确定出目标的运动参数。





目标跟踪 (Object tracking)





事件识别 (Event recognition)

事件的表示:



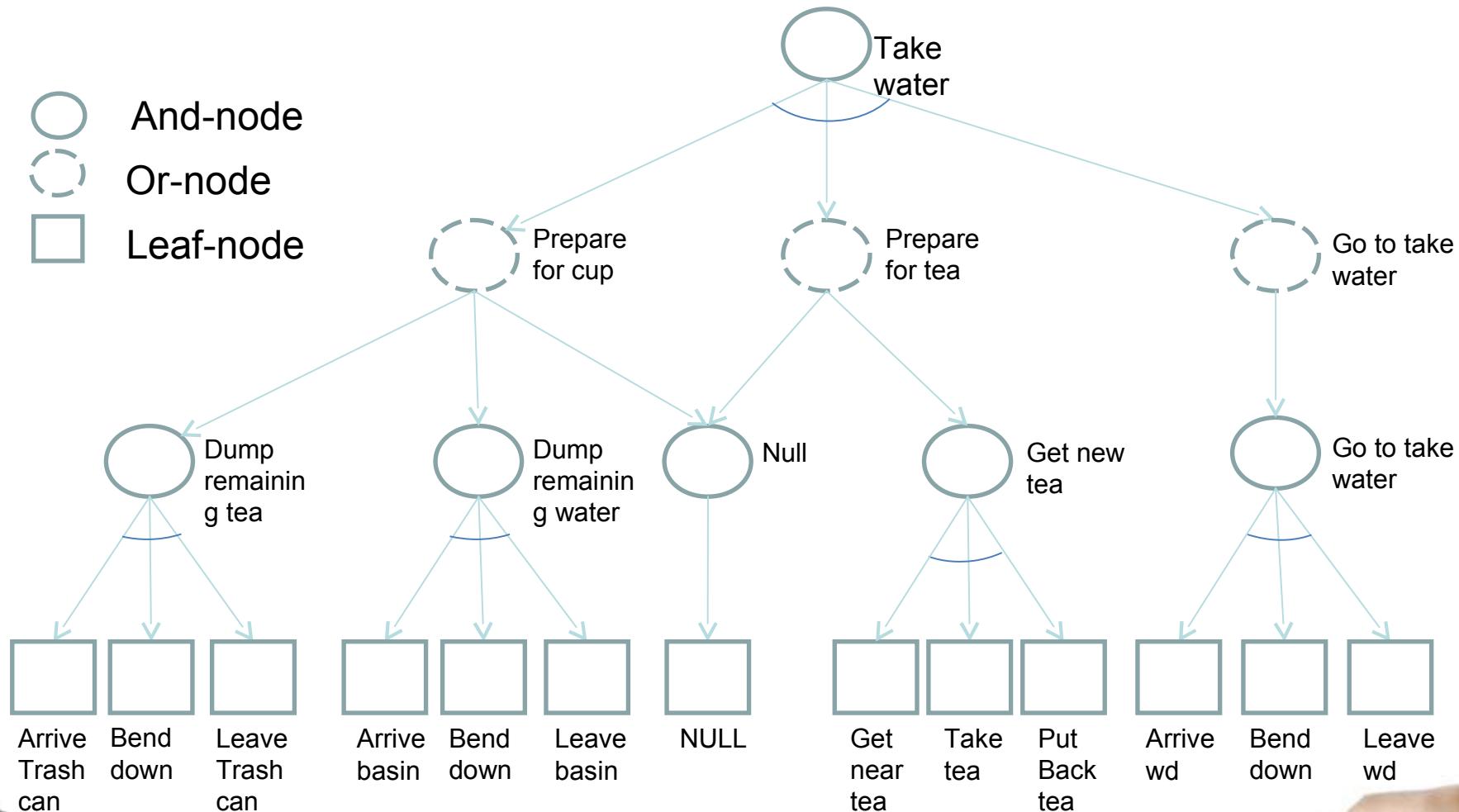
如何描述高层的语义事件?



事件识别 (Event recognition)

基于与或图的事件表示： Events Representation by AOG

- And-node
- Or-node
- Leaf-node





事件识别 (Event recognition)

Atomic Actions	Fluents	Symbols		Examples
		Foreground	Background	
Shake Hands(P1,P2)	Near(P1,P2) And Touch(P1.hand, P2.hand)			
Use Dispenser(P3)	Bend(P3) and Near(P3, A) And Touch(P3.hand, A)			
Pick up Phone(P4)	Touch(P4, B) And On(B)			

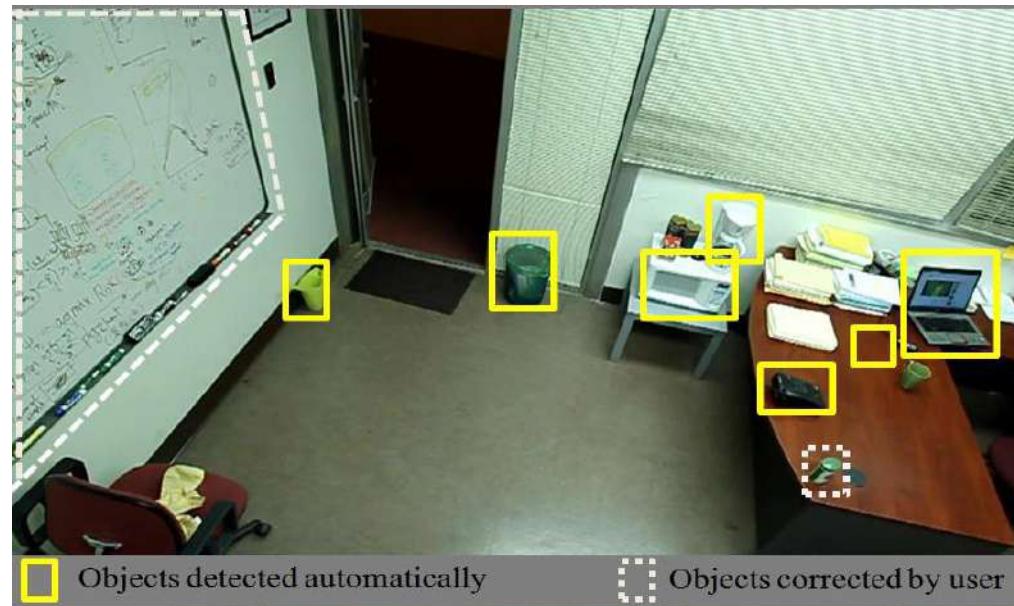
Examples of atomic actions, each relation is shown by 2 half-circles that are bonded together. For the action of 'shaking hands', □,□ are both agents.





事件识别 (Event recognition)

场景分割:



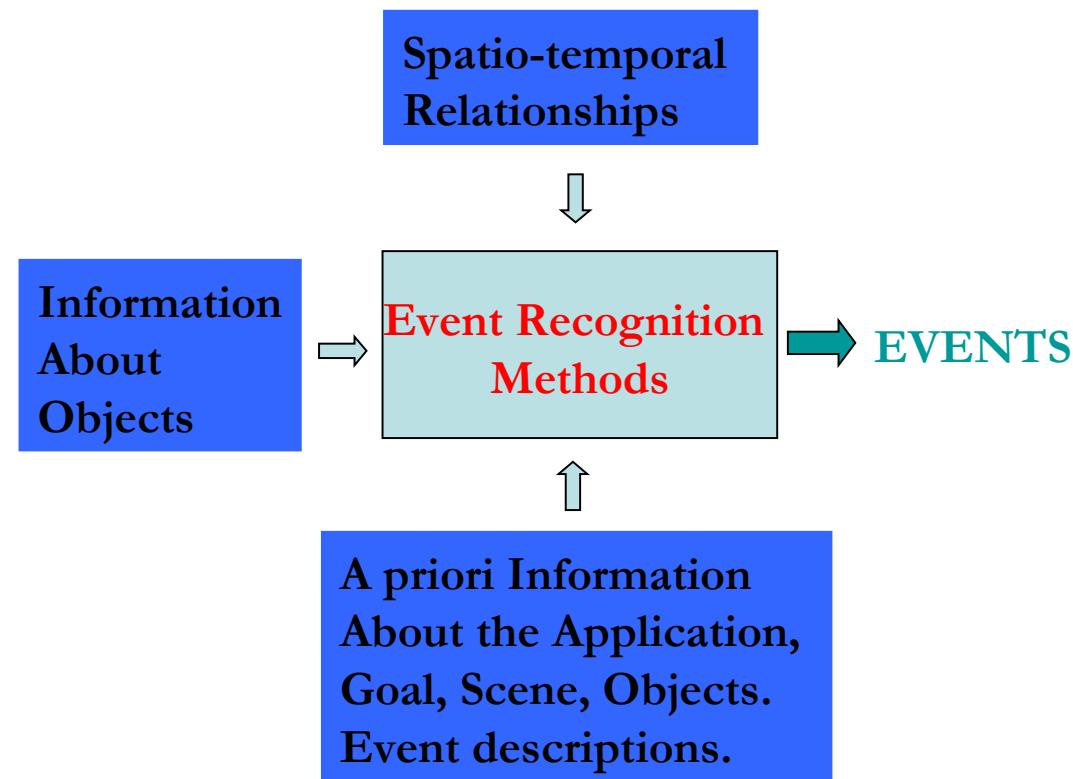
chair
desk top
computer
paper
phone
cup
tea box
microwave
water dispenser
trash can
basin
whiteboard
floor1
floor2
floor
wall





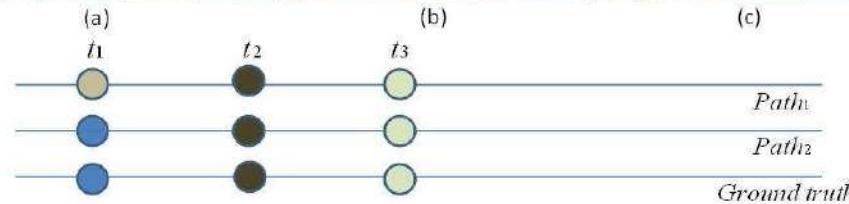
事件识别 (Event recognition)

事件识别 Event Recognition





事件识别 (Event recognition)

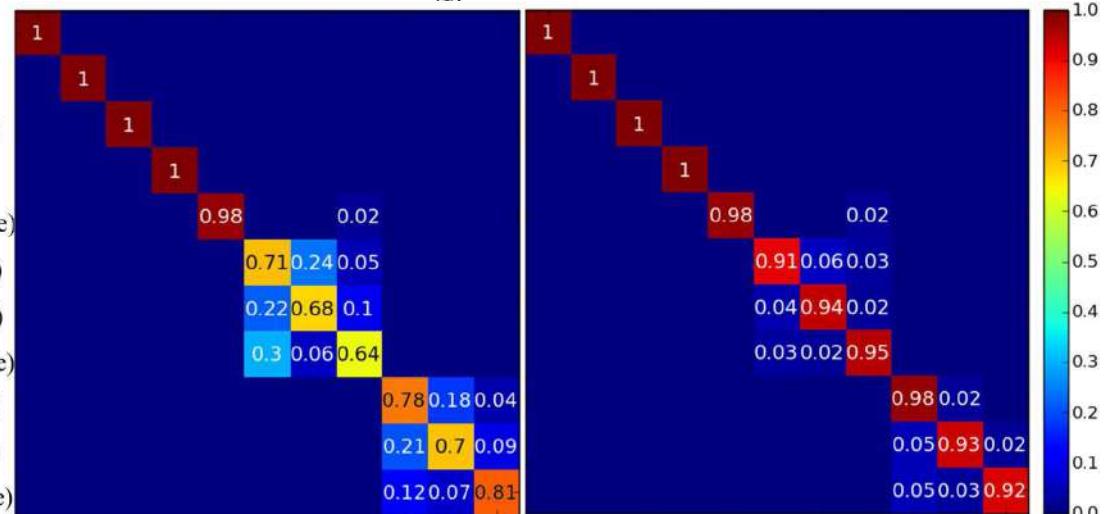


Atomic action:

● Arrive chair	● Arrive phone	● Sit down
● Leave chair	● Use phone	● Leave phone

(d)

Enter(agent)
Exit(agent)
BendDown(agent)
StandUp(agent)
Touch(agent,phone)
Arrive(agent,table)
Arrive(agent,chair)
Arrive(agent,phone)
Leave(agent,table)
Leave(agent,chair)
Leave(agent,phone)





事件识别 (Event recognition)

插入事件的识别：



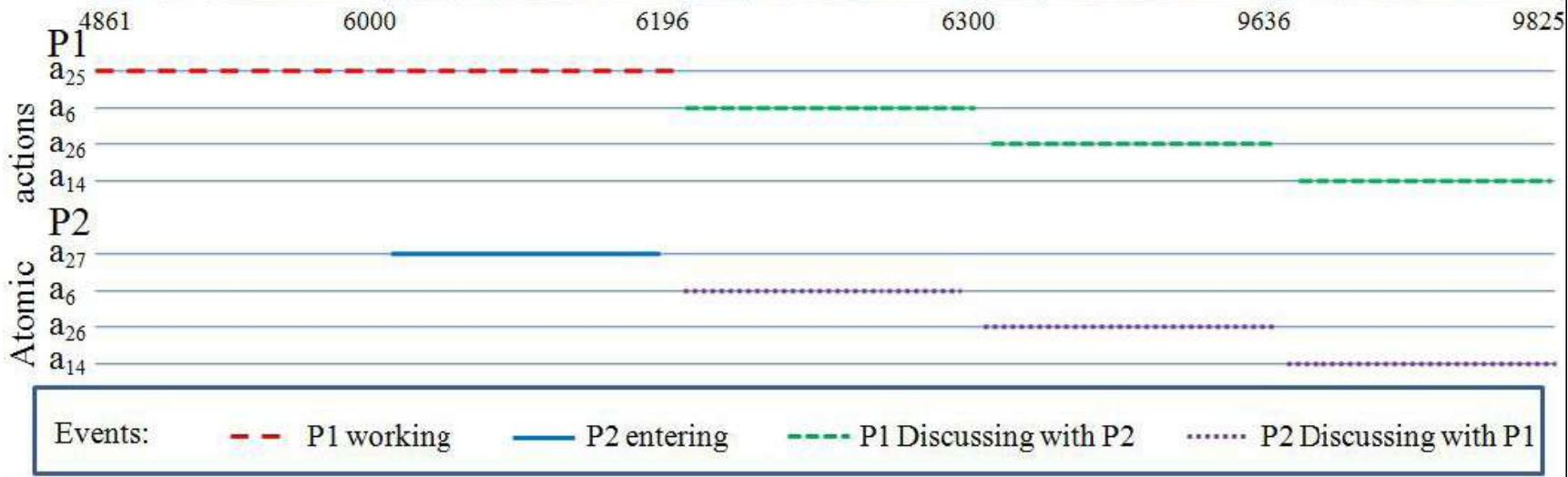
(a) Five judgement points of the sequential event of getting water and making a call



(b) Experiment results of event segmentation and event insertion



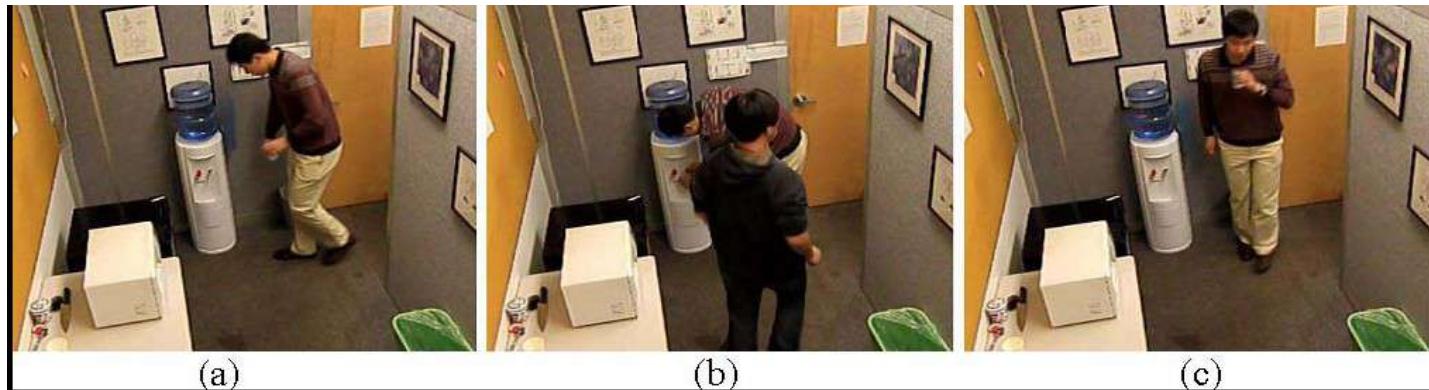
事件识别 (Event recognition)



Experiment results of event recognition which involve multiple agents.



事件识别 (Event recognition)



(a) (b) and (c) are three representative frame of atomic action 'arrive dispenser', 'use dispenser' and 'leave dispenser' respectively.



参考文献

- [1] Natarajan P and Nevatia R. Coupled hidden semi-markov models for activity recognition. *IEEE Workshop on Motion and Video Computing*, 2007.
- [2] Ivanov Y A and Bobick A F. Recognition of visual activitiesand interactions by stochastic parsing. *PAMI*, 8, 2000.
- [3] Pei M T, Jia Y D, and Zhu S C. Parsing video events with goal inference and intent prediction. *ICCV*, 2011.
- [4] Zhang Z, Huang K Q, Tan T N, and Wang L S. Trajectory series analysis based event rule induction for visual surveillance. *Proc. IEEE Int'l Conf. CVPR*, 2007.
- [5] Hakeem A and Shah M. Learning, Detection and representation of multi-agent events in videos. *Artificial Intelligence*, vol. 171, nos. 8-9, pp. 586-605, 2007.
- [6] Si Z Z, Pei M T, Yao B and Zhu S C. Unsupervised learning of event and-or grammar and semantics from video. *ICCV*, 2011.
- [7] Turaga P, Chellappa R, Subrahmanian V S and Udrea O. Machine recognition of human activities: A survey. *IEEE Trans. CSVT*, vol. 18, no. 11, pp. 1473-1488, Nov. 2008.
- [8]. Jiangen Zhang, [Wenze Hu](#), [Benjamin Yao](#), [Yongtian Wang](#), [Song-Chun Zhu](#): Inferring social roles in long timespan video sequence. *ICCV Workshops 2011*: 1456-1463
- [9] A. Litvin, J. Konrad, W. C. Karl. Probabilistic video stabilization using Kalman filtering and mosaicking. *IS&T/SPIE Symposium on Electronic Imaging, Image and Video Communications and Proc.*, Jan. 2003, Santa Clara, CA, USA
- [10] Carlos Morimoto, Rama Chellappa, Fast Electronic Digital Image Stabilization for Off-road Navigation, *IEEE Journal of Real-time Imaging*, 2, pp.285-296, 1996
- [11] S. Birchfield, [Elliptical Head Tracking Using Intensity Gradients and Color Histograms](#), *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Santa Barbara, California, pages 232-237, June 1998
- [12] T. Zhao, R. Nevatia, Tracking Multiple Humans in Crowded Environment, *Proc IEEE Conf on Computer Vision and Pattern Recognition (CVPR'04)*



数字图像处理

(Digital Image Processing)

山东科技大学



第8章

数字图像技术 应用系统设计

山东科技大学



8.5 医学图像处理

Medical Image Processing

山东科技大学
图像处理课题组



8.5.1 医学图像处理简介

8.5.1.1 医学图像处理

医学图像处理是利用数学的方法和计算机这一现代化的信息处理工具，对由不同的医学影像设备产生的图像按照实际需要进行处理和加工的技术。

医学图像处理学科致力于对人类疾病的发现、诊断、辅助治疗，致力于为探索生命现象提供高水平的科学方法和工程技术手段。

8.5.1 节图片内容选自以下网站教学资源

<http://cc.usst.edu.cn/able.acc2.web/template/view.aspx?wmz=4074&coursetype=0&courseid=275&topmenuid=48976&menutype=4&contentid=4074&action=view&type=&name=&menuid=000002>



医学图像的类型

- ◆ X 射线 (X-ray) 图像
- ◆ CT (Computerized Tomography) 图像
- ◆ MRI (Magnetic Resonance Imaging) 图像
- ◆ 超声 (Ultrasonic) 图像
- ◆ PET (Positron emission tomography) 图像
- ◆ SPECT (Single Photon Emission Computed Tomography)
图像



8.5.1.2 医学图像处理技术的主要内容

医学图像处理涉及的主要内容包括：医学影像的数字化基础、医学图像的运算、医学图像变换技术、医学图像增强技术、医学图像分割技术、医学图像配准、医学图像三维可视化等。

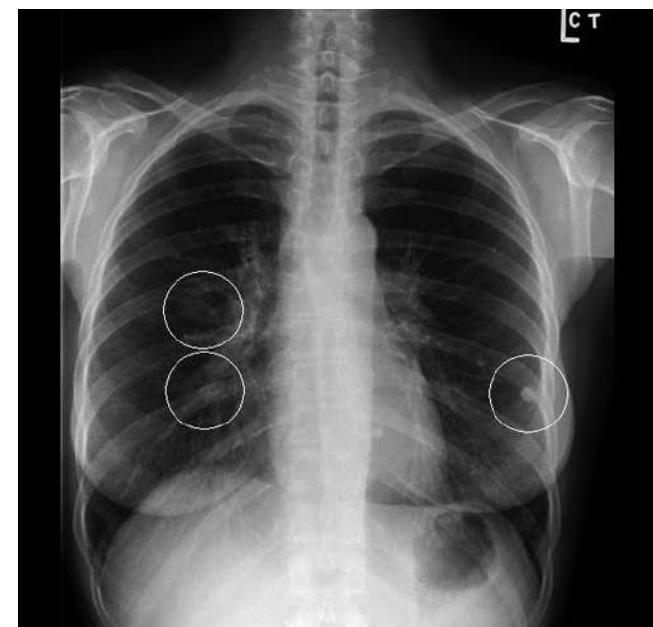
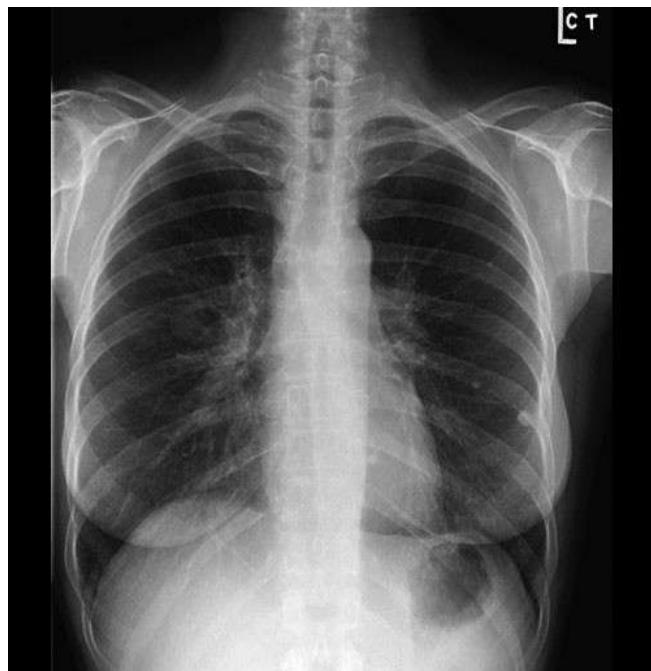


8.5.1.3 医学图像处理技术的主要应用领域

- 定量影像学（精确诊断、精确治疗）
- 基于医学图像的计算机辅助诊断 / 检测；
- 虚拟内窥镜；
- 外科手术术前计划系统；
- 影像引导下的放射治疗系统；
- PACS 系统 / 远程诊疗
- ...

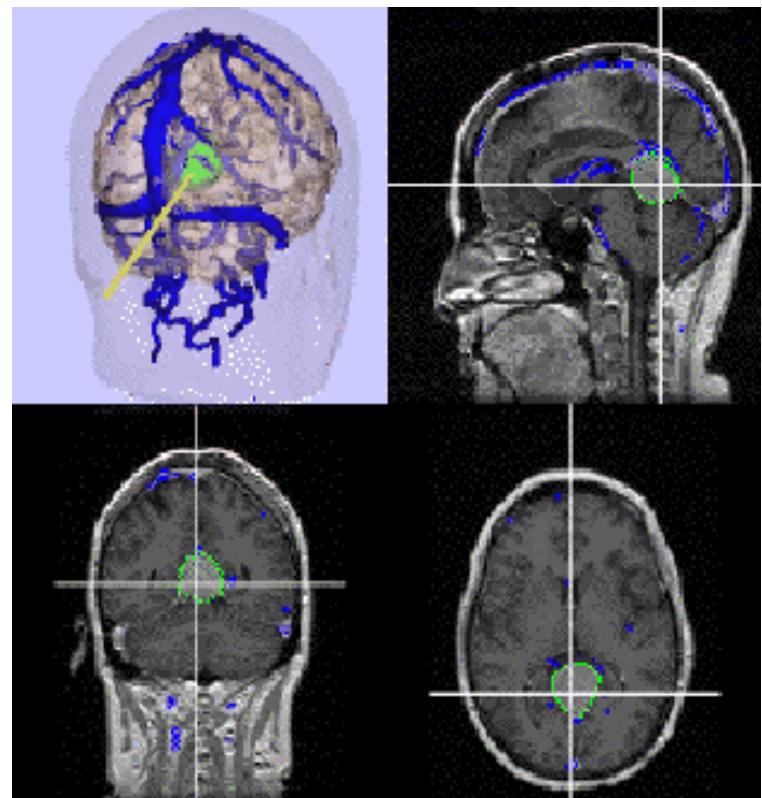
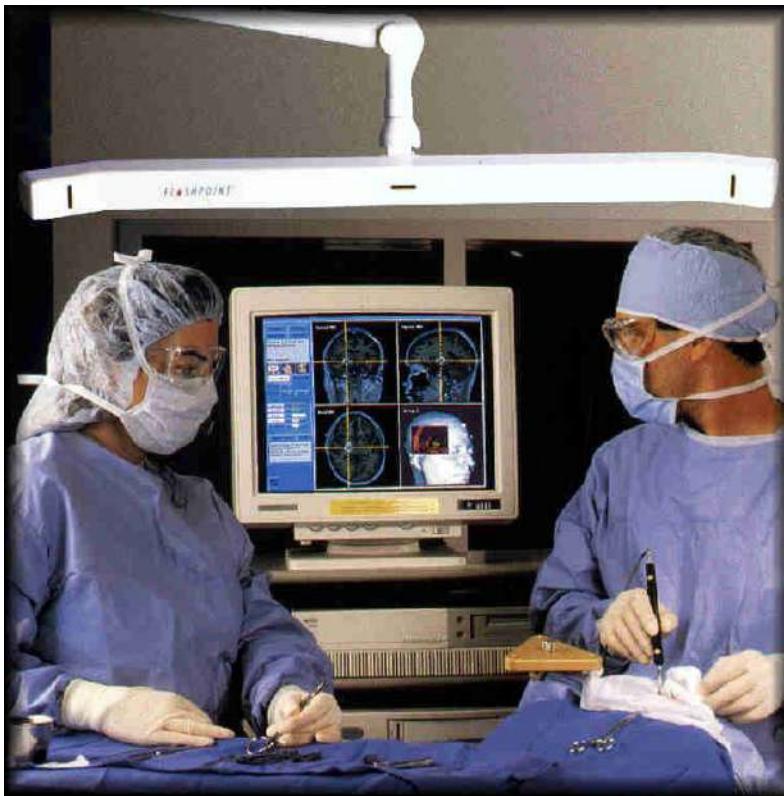


计算机辅助诊断



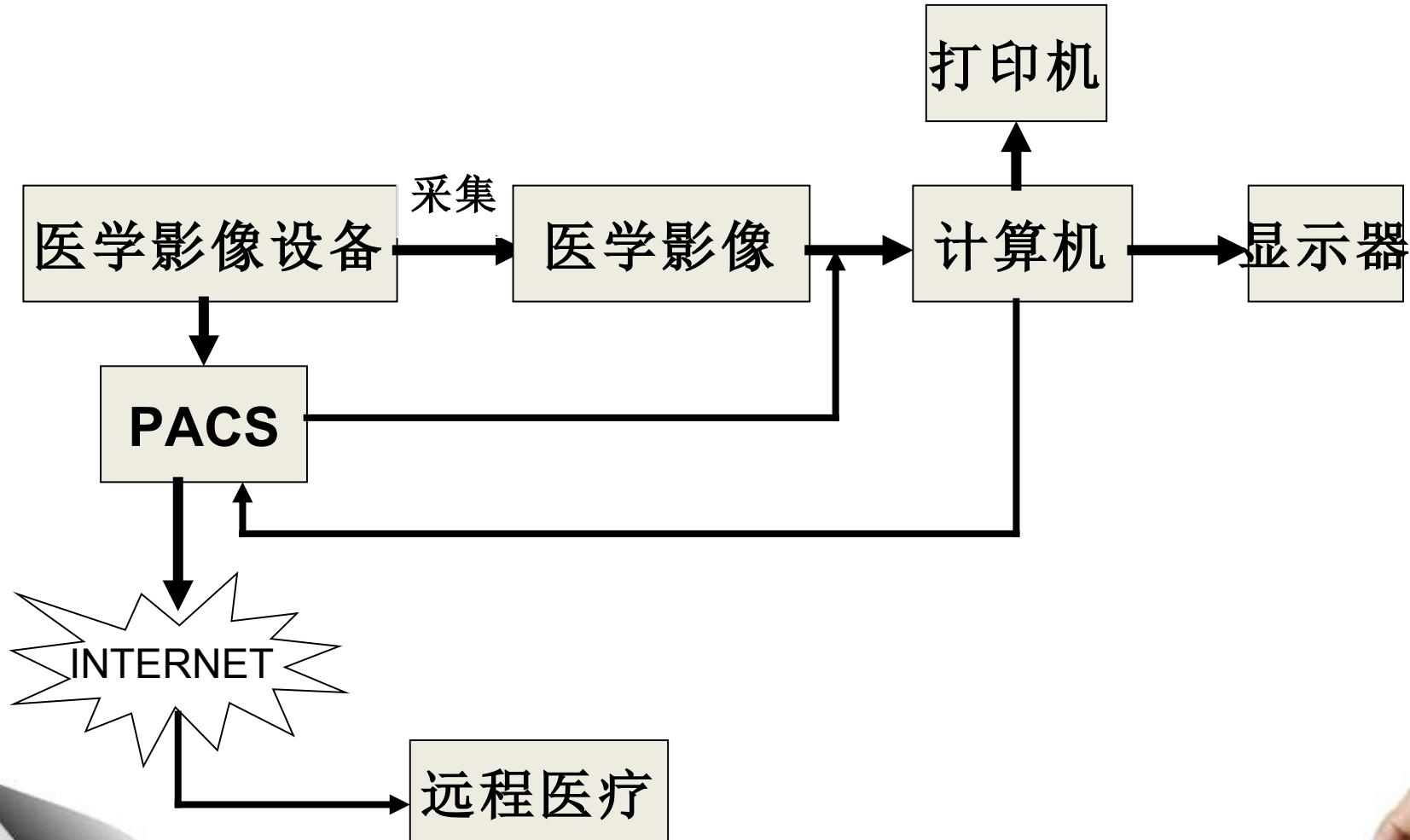


基于图像导航的手术与治疗





8. 5. 1. 4 医学图像处理的基本过程





8.5.2 医学图像获取 --- 医学影像技术

1、医学成像技术概述

2、X线成像

3、超声成像

4、CT成像

5、核医学成像

6、MRI/FMRI 成像

7、医学图像的发展趋势

8.5.2 节图片内容选自以下网站教学资源

<http://cc.usst.edu.cn/able.acc2.web/template/view.aspx?wmz=4074&course type=0&courseid=275&topmenuid=48976&menutype=4&contentid=4074&action=view&type=&name=&menuid=000002>



8 . 5 . 2 . 1 医学影像技术的发展

德国科学家伦琴于 1895 年发现的 X 射线并由此产生的 X 线成像技术，第一次无损地为人类提供了人体内部器官组织的解剖形态照片，由此引发了医学诊断技术的一场革命。

形成了放射诊断学 (diagnostic radiology) 的新学科。

50 年代到 60 年代，超声成像和 γ 闪烁成像 (γ -scintigraphy)。

70 年代和 80 年，X 线计算机体层成像 (CT)、磁共振成像 (MRI) 和发射体层成像 (Emission Computed Tomography, ECT)，如单光子发射体层成像 SPECT 与正电子发射体层成像 PET 等新的成像技术。



CT 已从早期的单纯的头颅 CT 发展为超高速多排螺旋 CT、电子束 CT。在速度提高的同时，扫描最薄层厚也从早期的 10mm 到现在的 0.5mm 以下，图像分辨率也达到了 1024*1024。

CT 的应用不仅在于早期横断面成像，同时可以作细腻的三维重建，虚拟内窥镜，手术立体定向，CT 血管成像（Computed Tomography Angiography，CTA）等。

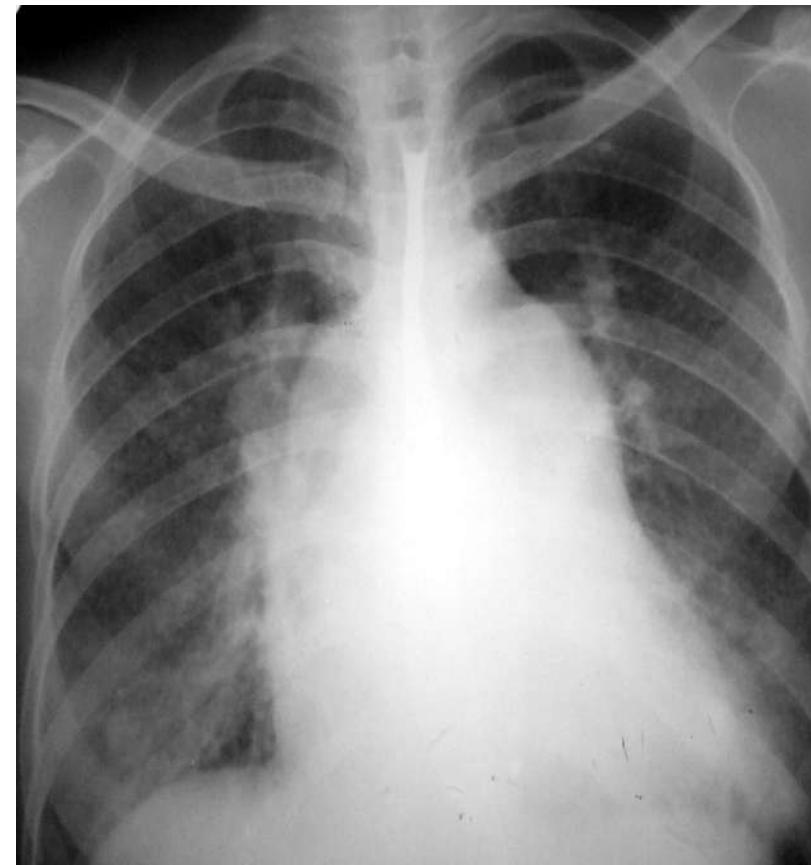
MRI 也从早期的永磁体、低场强发展到现在的超导、高场强，分辨率在常规扫描时间下提高了数千倍，磁共振血管成像（Magnetic Resonance Angiography，MRA）已成为常规检查项目，同时磁共振功能成像以及磁共振波谱（MRS）技术正在迅速发展之中



8 . 5 . 2 . 2 X 线成像



1895 年伦琴发现 X 线，通过 X 线第一次观察到 人体内部的结构，为医生确诊疾病的病因提供了重要的信息。



伦琴因发现 X 射线获得首届诺贝尔物理学奖（1901年）



X 射线成像（传统 X 线成像）

原理：X 线束直接对准病人需检查的部位，从病人后面透射过人体，借助荧光屏产生可见光并使胶片感光（胶片屏幕组合）。

照片上某个像素的亮度反映穿过人体到达胶片的 X 射线的强度，它与人体对 X 射线的吸收量成反比。

组织密度、厚度不同，对 X 射线吸收、透射 X 射线强度不同，如肺组织对 X 线的衰减效应比骨骼小得多。

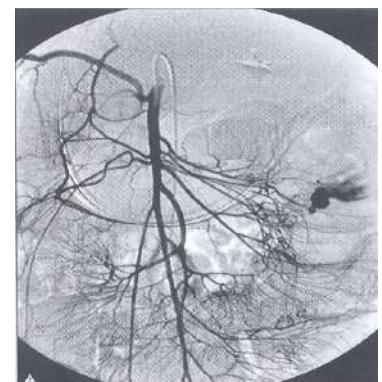
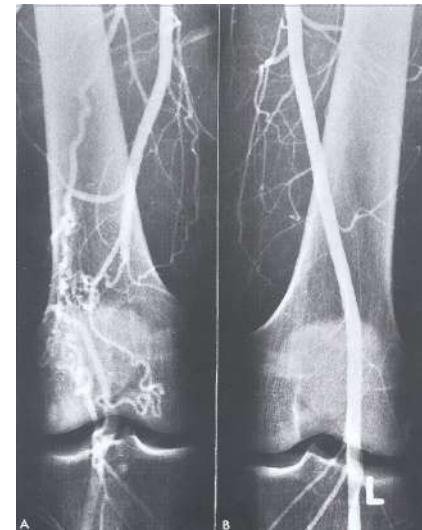
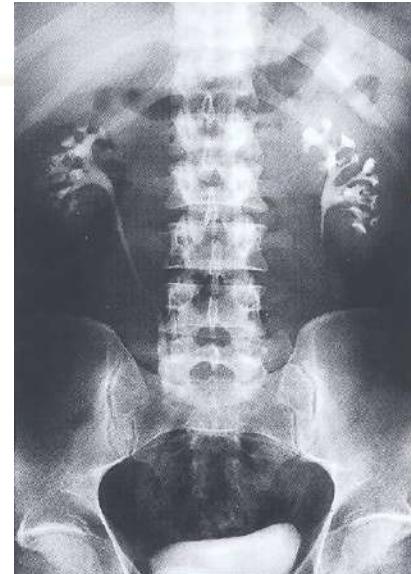
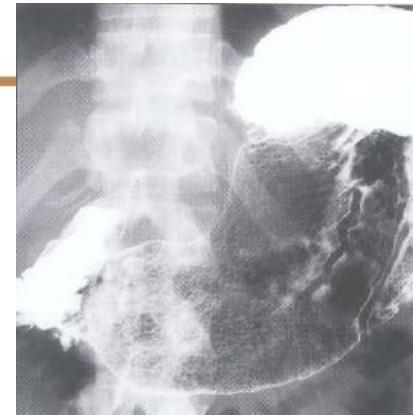
缺点：不能反映组织和病灶的三维空间；不是数字化的形式





经过百年的发展，应用 X 线机可观察人体内部的骨骼、肺结核病变等，通过造影技术，可以观察心脏、血管及消化道等管状器官。

随着相关科技的发展，X 线投影成像技术，一直在不断地改进和发展。

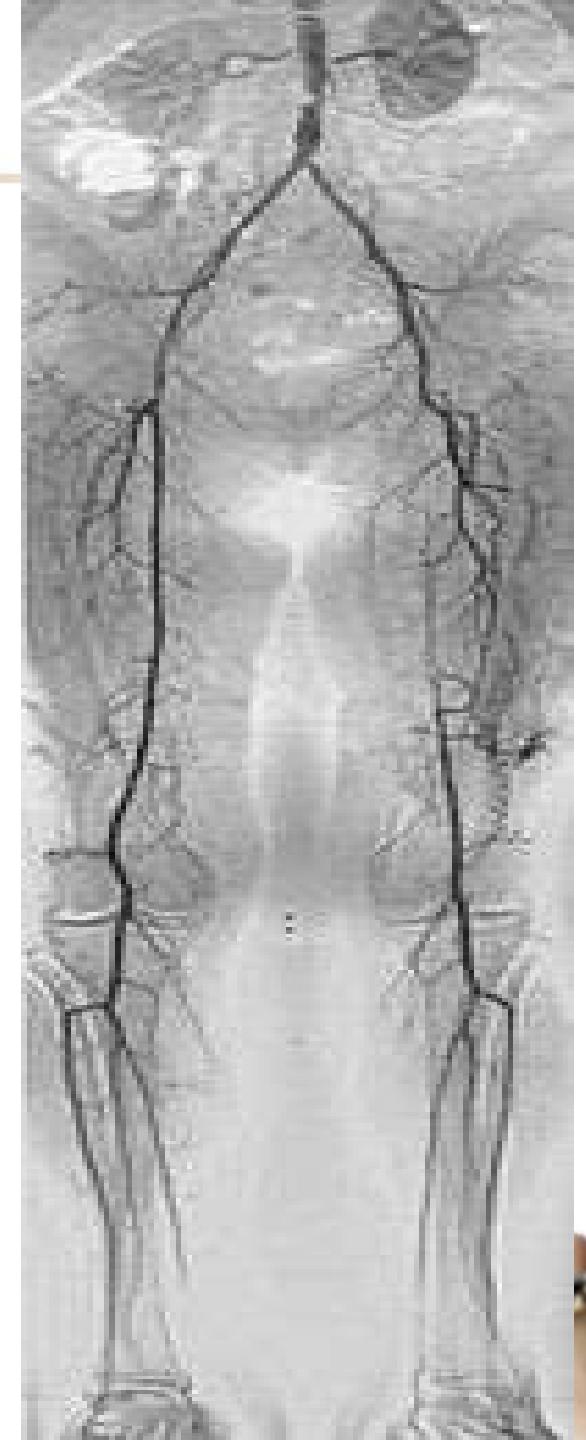


X 线成像技术没有从根本上改变 3D 物体投影到 2D 平面的问题。由于处于不同深度的器官的像是重叠在一起的，从而使得 X 线影像的空间分辨率受到一定限制。



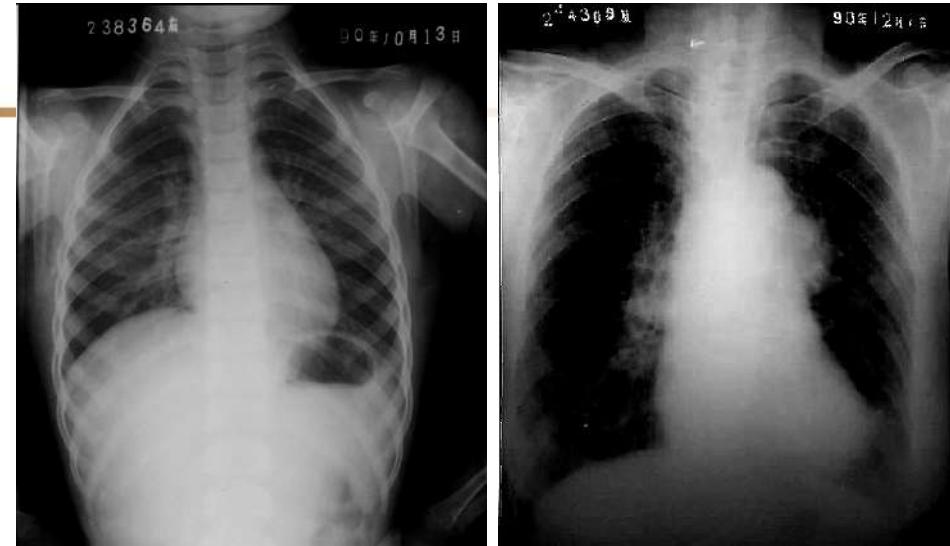
数字减影血管造影（Digital Subtraction Angiography，DSA）是利用数字图像处理技术中的图像几何运算功能，将造影剂注入前后的数字化X线图像进行相减操作，获得两帧图像的差异部分——被造影剂充盈的血管图像。

目前 DSA 有时间减影（temporal subtraction）、能量减影（energy subtraction）、混合减影（hybrid Subtraction）和数字体层摄影减影（digital tomography subtraction）等类型。





计算机 X 线摄影 (computed radiography, CR) 是 X 线平片数字化的比较成熟的技
术。

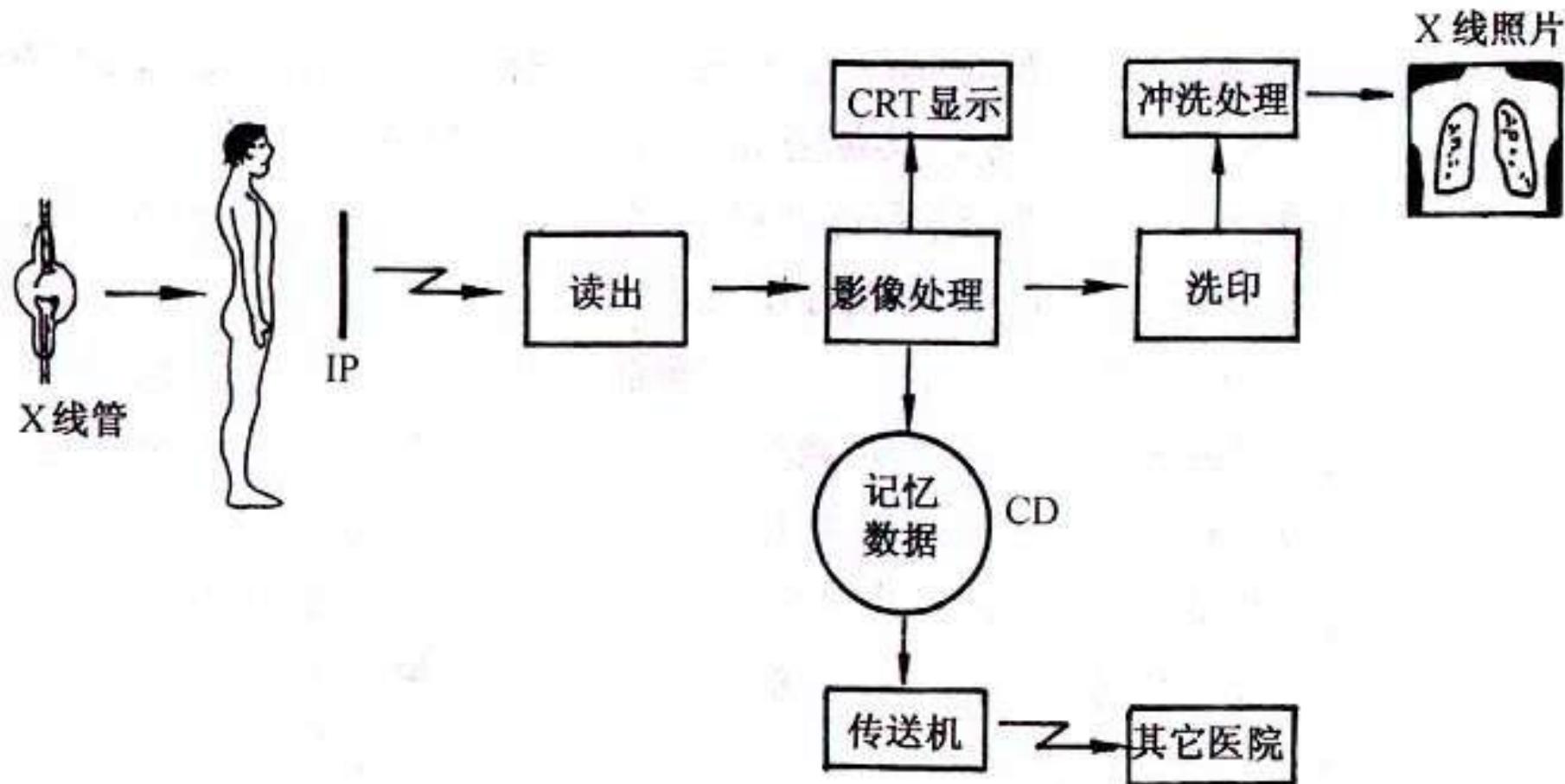


CR 系统是使用可记录并由激光读出 X 线成像信息的成像板 (Imaging Plate , IP) 作为载体，经 X 线曝光及信息读出处理，形成数字式平片图像。





计算机 X 线摄影 , CR

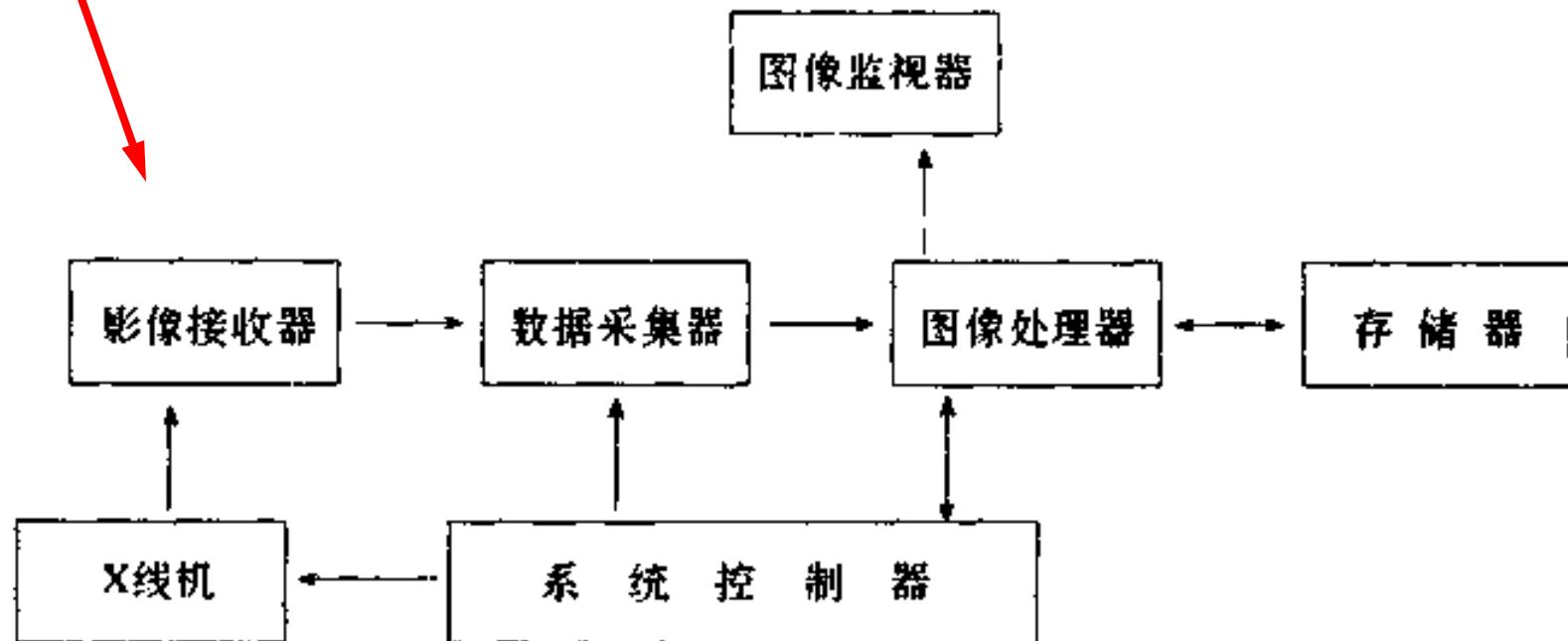


CR 系统示意图



•Digital Radiography, DR

(Flat Panel Detector,
FPD)





8 . 5 . 2 . 3 超声成像

X 线对人体健康是有害的，在第二次世界大战时期发展起来的雷达和声纳的基础上，应用超声脉冲反射原理发展了各种超声成像技术。

超声可以探查出非常细微的病变组织，是 X 线摄影的有力补充。超声成像也是除了 X 线以外使用最为广泛的医学成像工具。超声成像依据的是脉冲 - 回波技术，这个技术和雷达技术相似。

超声成像的缺点是图像对比度差，图像的重复性依赖于操作人员。另外，超声检查的视野有限，难以显示正常组织及较大病变的全貌，也不利于与其它检查图像（如 CT ， MRI ）进行对比。



超声图像

原理：利用超声在人体组织内的回波原理，接收回波成像。



16周胎儿的超声图像



肝脏

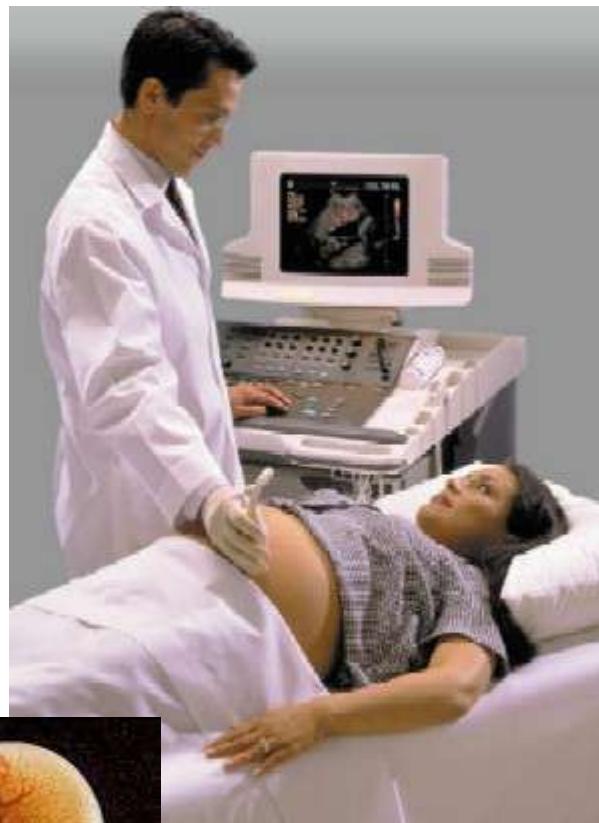
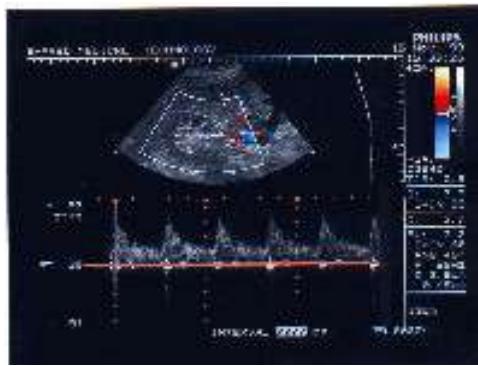
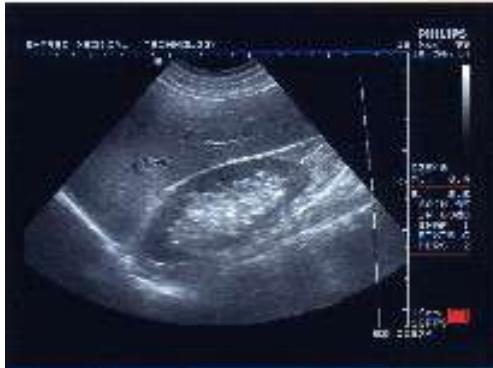


肾脏

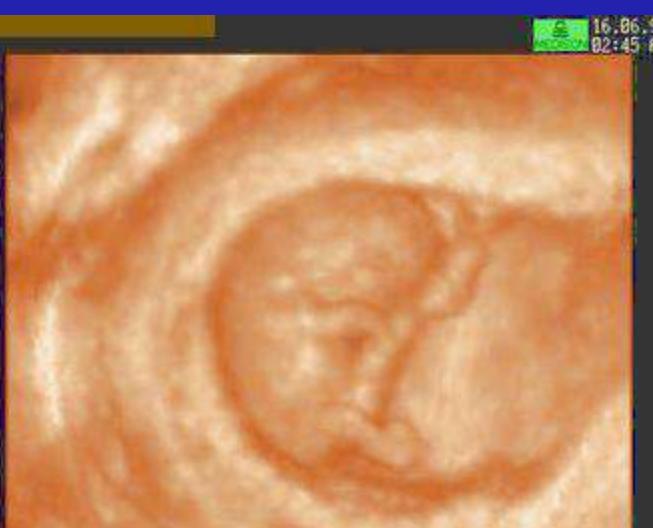


颈动脉

超声图像



超声图像





成像

- 随着计算机技术的发展，1972年出现了一场医学成像技术的革命。英国工程师 Hounsfield 因研制成功第一台头部扫描 CT，并于 1979 年获得了诺贝尔医学生物学奖。这是在诺贝尔奖的历史上第一次由工程技术人员获奖。



X-CT 计算机断层图像

传统 X 线图像是投影图像，不能显示真实的器官几何形态，在三维空间前后排列的器官将投射到一个叠加的二维图像上。为了得到一个三维图像，必须获得大量不同角度的图像，通过这些图像，放射学家和医生可构成他们感兴趣的三维图像。

原理：X 线管和探测器都装在支架上，可沿一直线移动（平移）和旋转，线管产生已知强度的细束射线，探测器则测量透过的射线强度，得到组织断层图像或三维图像



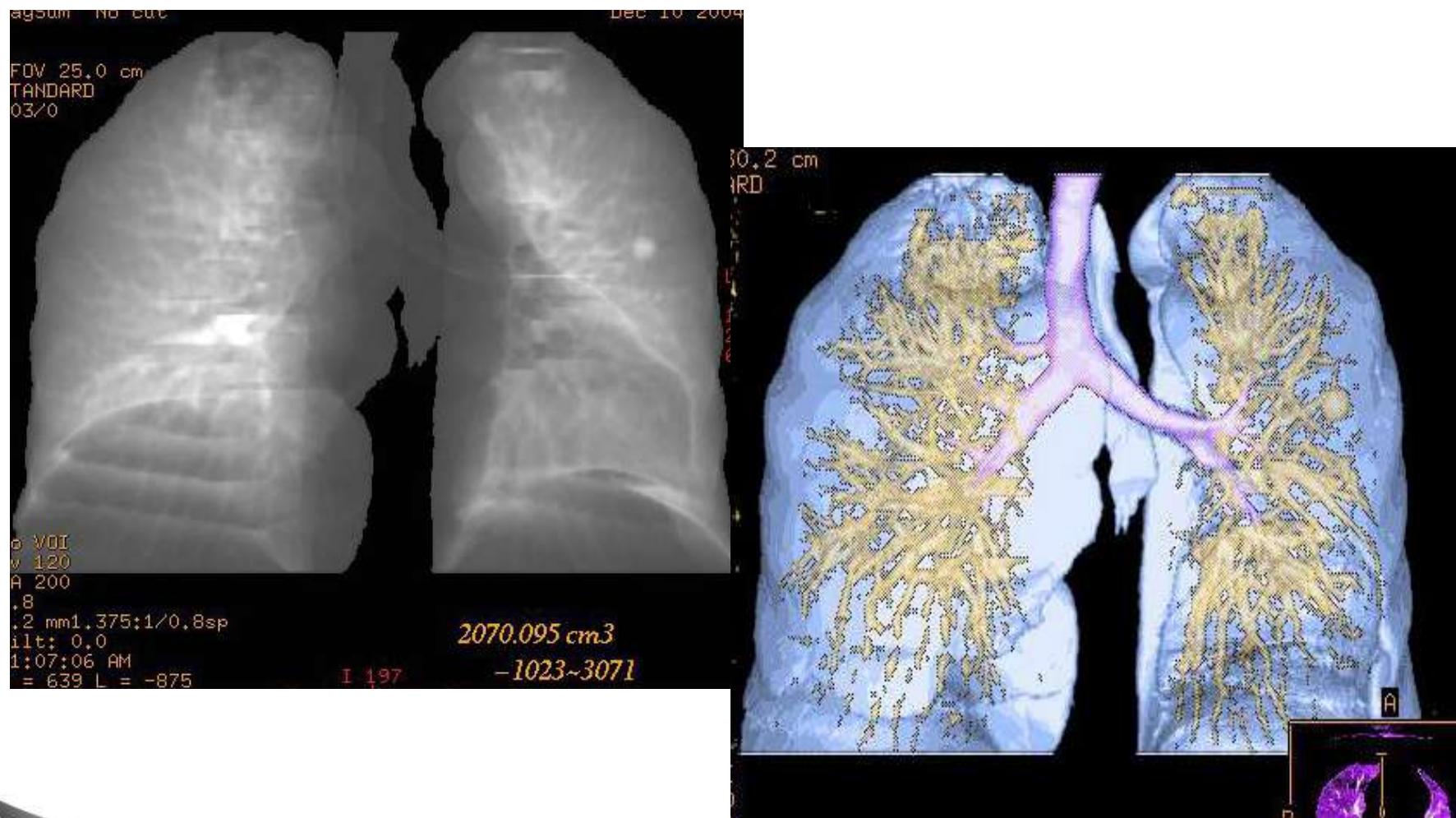


与传统的 X 线检查手段相比，CT 具有以下优点：能获得真正的断面图像，具有非常高的密度分辨率，可准确测量各组织的 X 线吸收衰减值，并通过各种计算进行定量分析。

CT 成像解决了传统 X 线成像因组织重叠造成的图像分辨率不高的问题，实现了组织器官的断层解剖结构的成像。但是，由于与 X 线成像技术一样，CT 成像也是通过检测人体对 X 射线的吸收量而获得的图像，因此，CT 成像对软组织获得的图像的密度分辨率远没有 MRI 高。

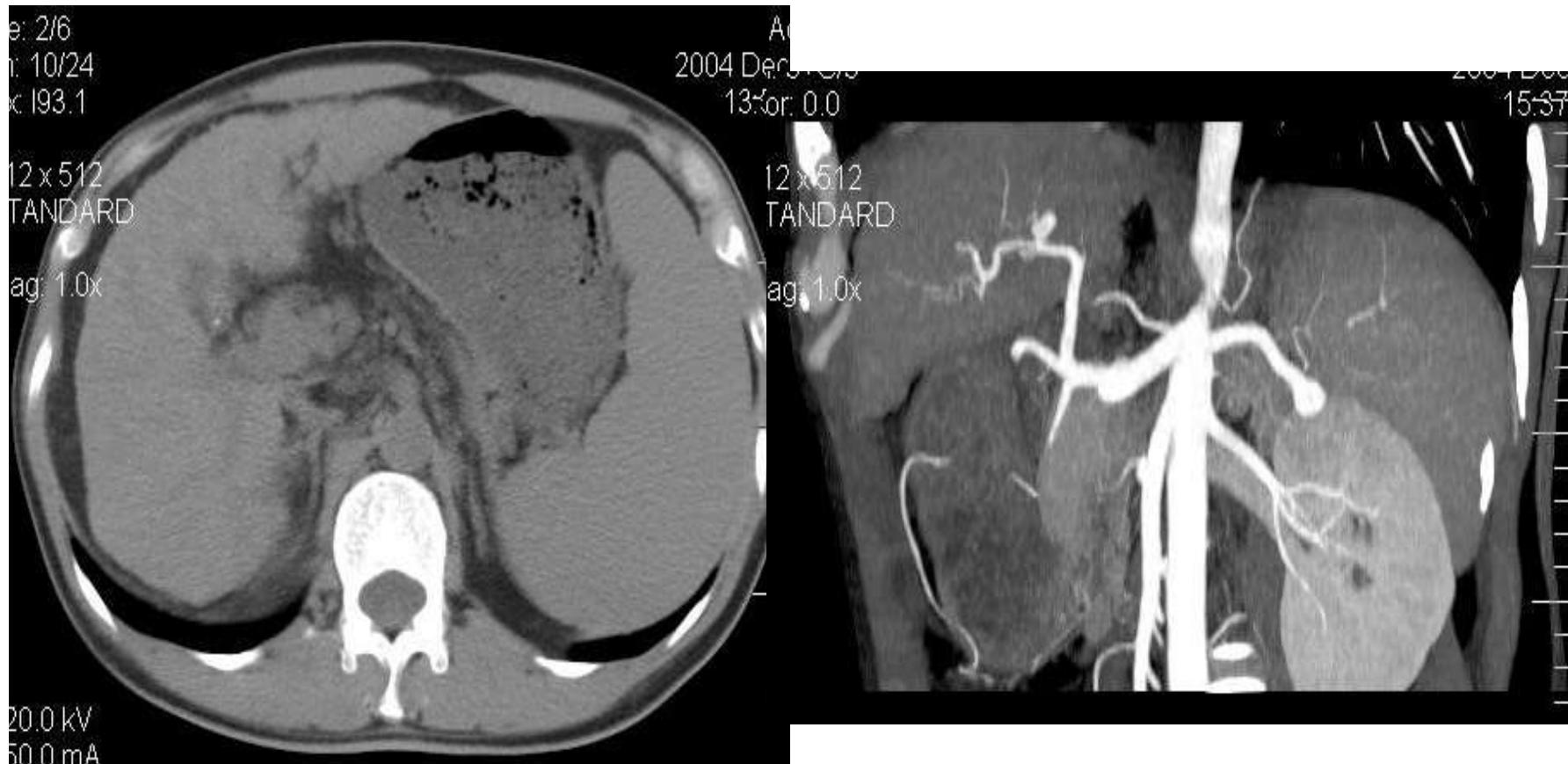


CT 图像





CT 图像





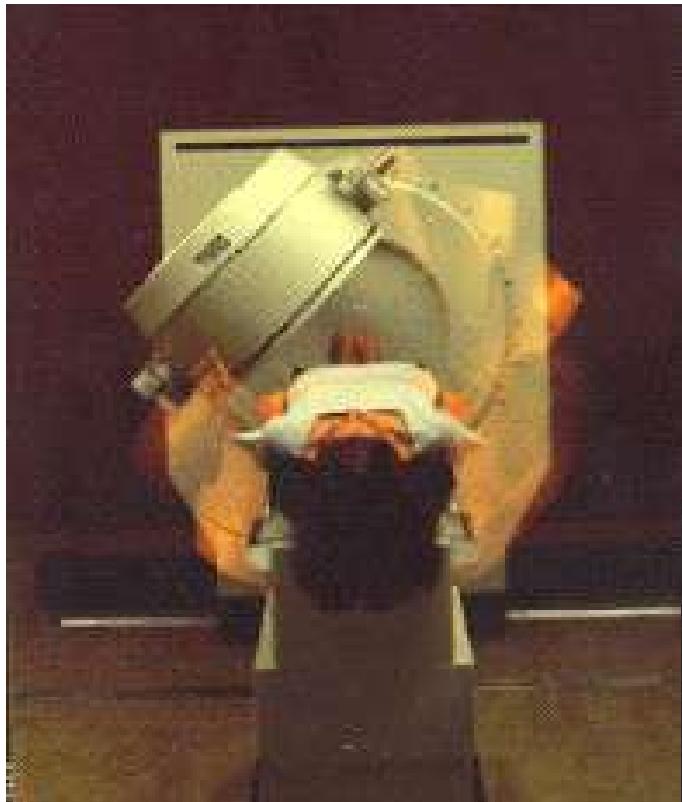
8.5.2.5 核医学成像

尽管 CT 技术解决了 X 线成像技术所不能解决的断层成像问题，但是 CT 也有其局限性，如对血管病变，消化道腔内病变以及某些病变的难以定性等，而且只能对组织结构进行成像，不能进行功能性的影像诊断方法如超声、CT、MRI 等提供的是人体解剖学变化的信息。而核医学成像提供的是人体组织器官的新陈代谢变化方面的信息，既功能性信息，这些信息有助于更早地发现疾病，并判断疾病的性质及发展程度。

目前，在核医学领域广泛使用的影像技术是 PET 和 SPECT，这两种成像技术又统称为发射型计算机断层 ECT（Emission Computed Tomography）。

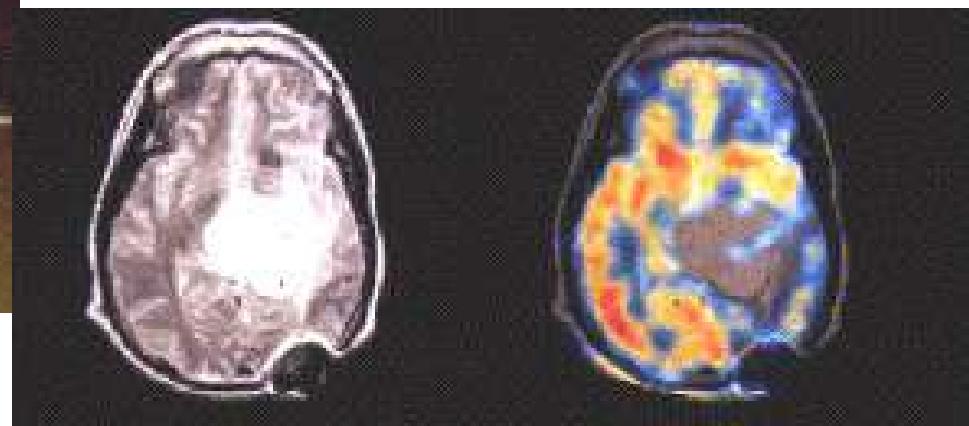


SPECT(单光子发射 CT) 成像



伽马照相机旋转

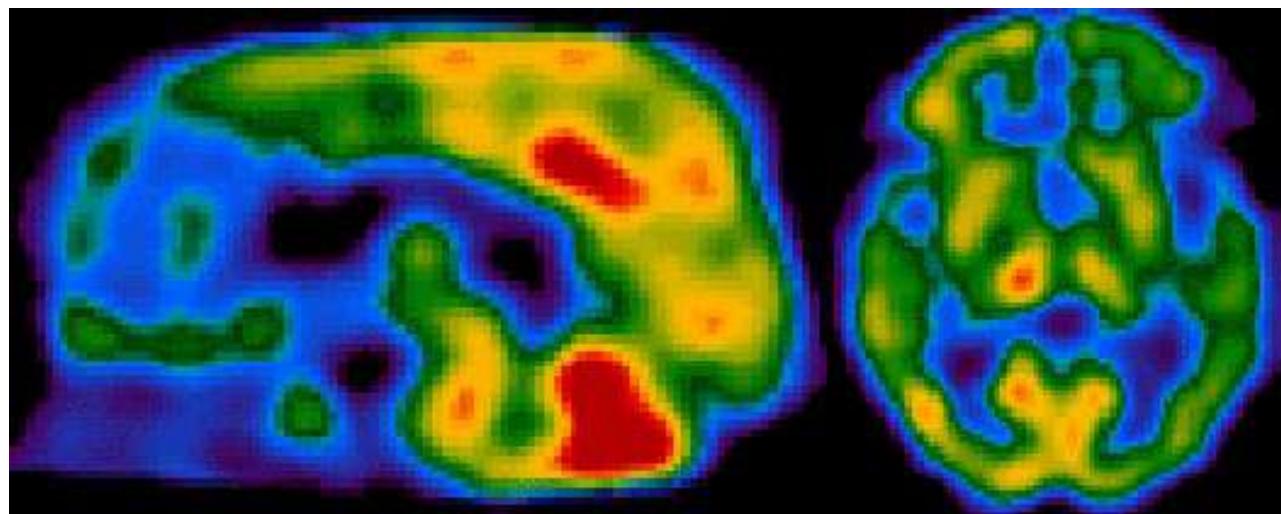
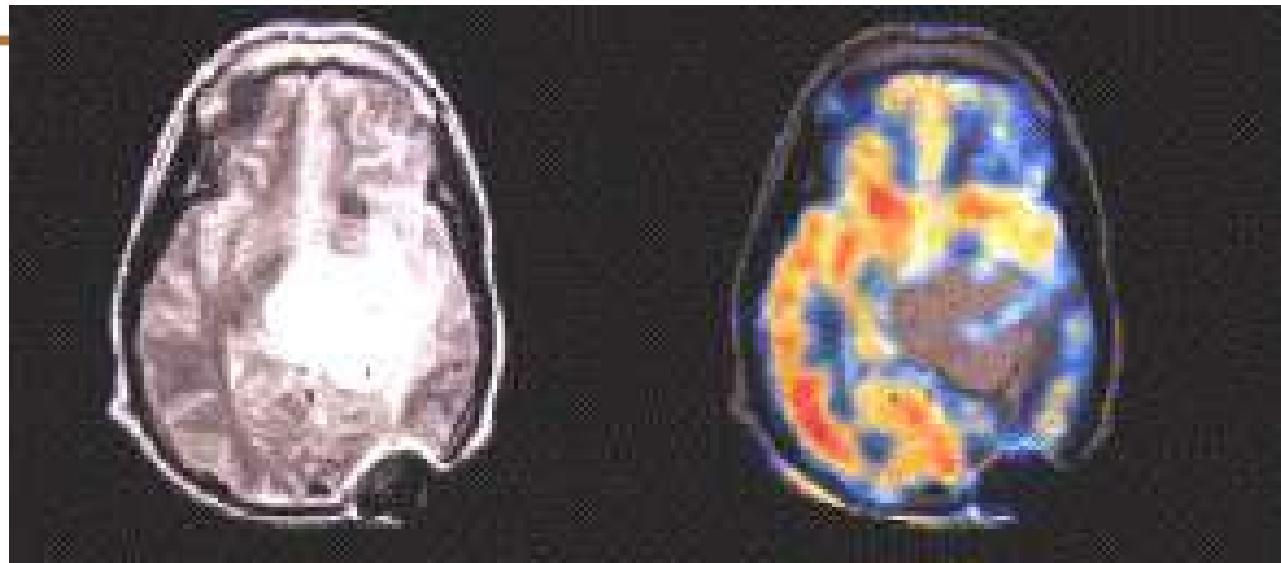
原理：
放射性同位素发射伽马射线，经人体后射线衰减，射线透过晶体转换成荧光光子，设备接受光子形成图像。



脑部特性

SPECT 可用来处理心脏的三维重建，也可用于脑和骨骼的研究等等







8.5.2.6 PET 图像

PET 又称正电子发射计算机断层显像仪 (Positron Emission Tomography , PET) , 是继 CT 之后出现的又一断层扫描技术 , 是核医学和 CT 技术相结合的产物 , 代表了核医学发展的最高水平 , 于 1976 年应用于临床。

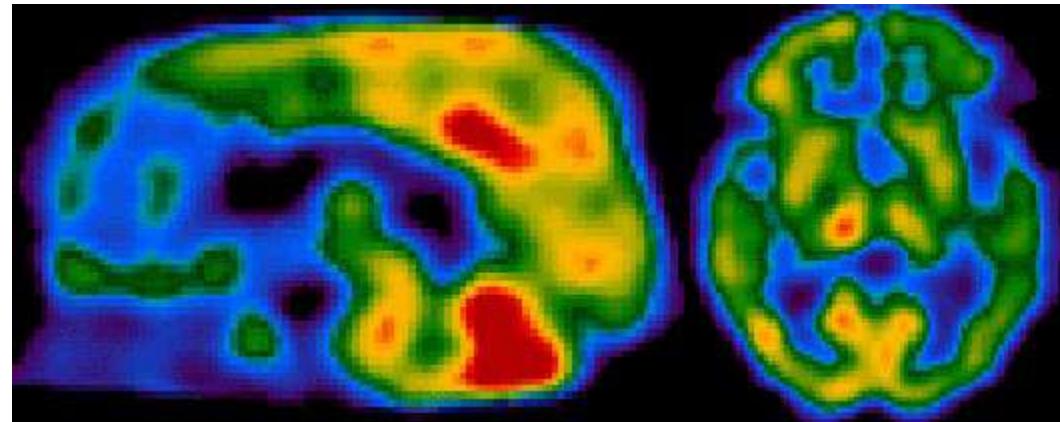


PET 成像

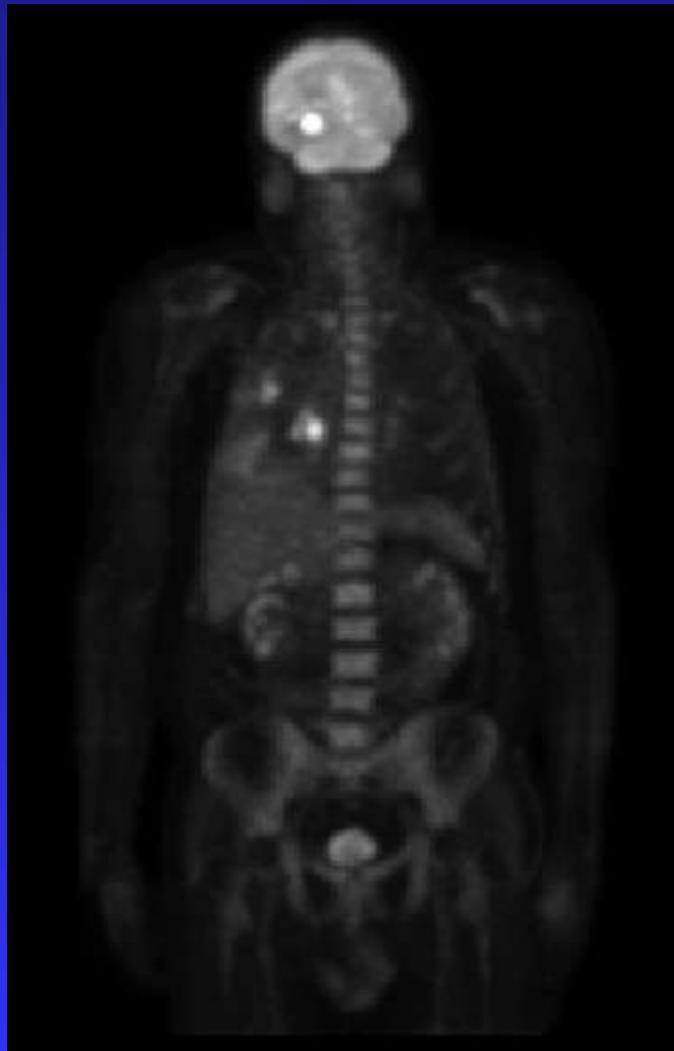
原理：

放射性同位素注入人体，同位素的正电子在泯灭时发射伽马射线，经检测器阵列接收，根据接收强度成像。

它反映活体靶组织在某一时刻的血流灌注、糖／氨基酸／核酸／氧代谢或受体的分布及其活性状况，可同时给出相应的活性生理功能参数

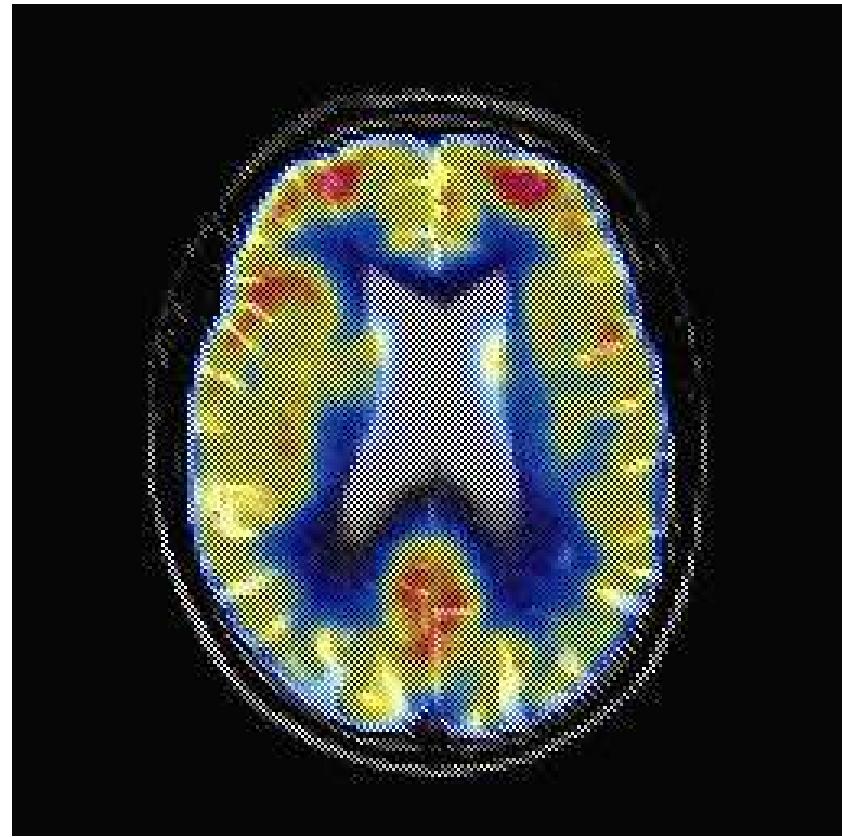
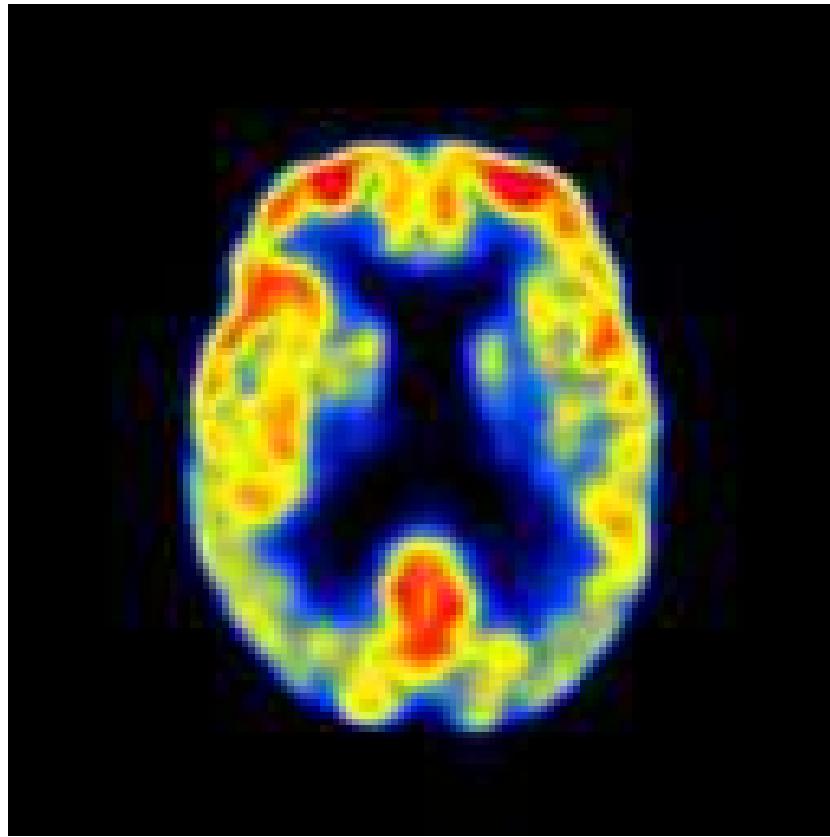


PET



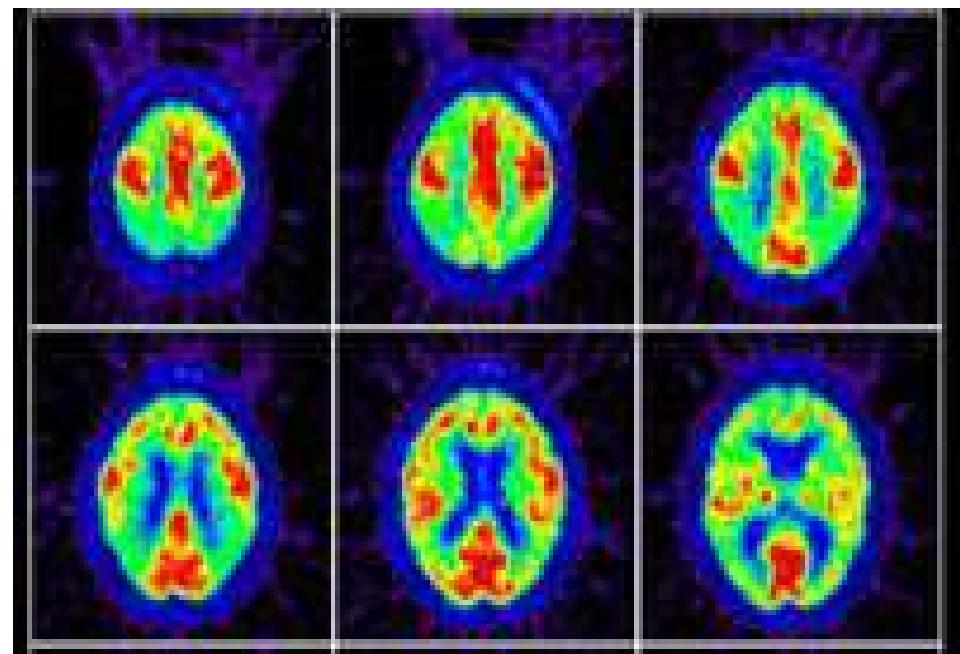
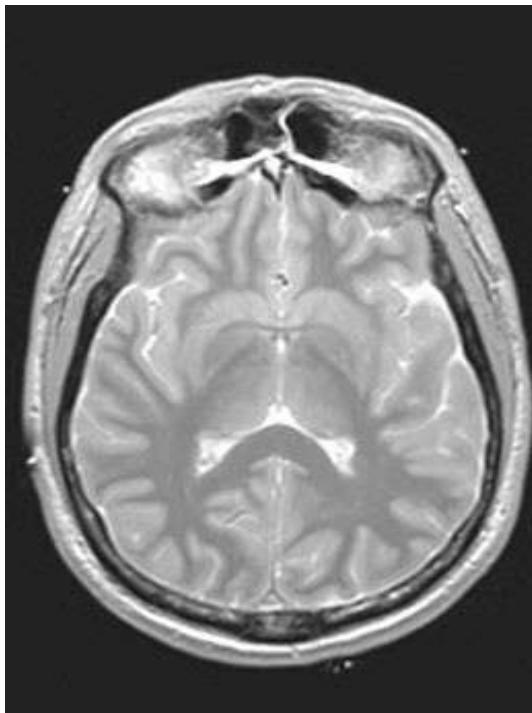


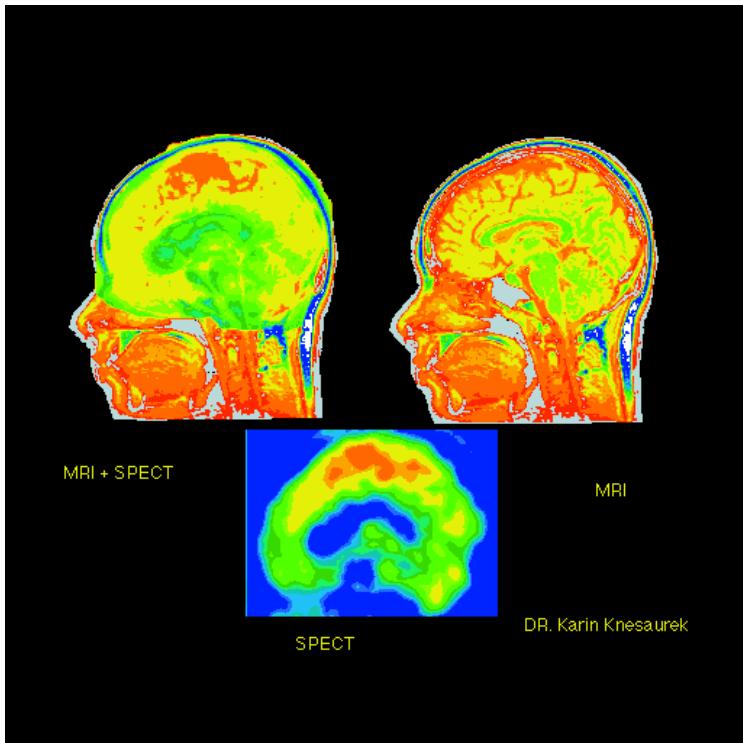
PET/MRI 融合图像





MRI 图像与 PET 图像比较





解剖结构成像（如 CT 图像）具有很高的空间分辨率，而核医学成像能得到不同器官或脏器的功能性信息。近年来，人们利用图像融合 (fusion) 技术，把解剖结构图像（CT 或 MRI 图像）与核医学图像融合在一起，使得在一幅图像上既包含组织结构的信息又包含组织功能的信息。



8. 5. 2. 7MRI / fMRI 成像

磁共振成像 (Magnetic Resonance Imaging, MRI) 是在 X-CT 出现后又一重大发明，它是医学、化学、物理学、电子学、电子计算机技术、X-CT 技术和磁共振波谱技术等学科相结合的产物。MRI 所提供的信息比 X-CT 多。因此，有人说 MRI 是二十一世纪的影像技术。



MRI 磁共振图像

原理：是以某种方法如通过测量断层组织氢核的密度，显示身体中某一断面的组织分布图像，由此而获得身体中某断面的二维图像。



大脑



膝盖



靠近肾脏的动脉

特点：显像时间快、图像清晰



MRI 的物理学基础是核磁共振 (nuclear magnetic resonance , NMR) 现象，其本质是一种能级间跃迁的量子效应。

NMR 现象是由于人体中的原子核吸收了来自外界的电磁波后，产生了共振现象，然后原子核再将其所吸收的能量以电磁波形式释放出来产生图像的过程。

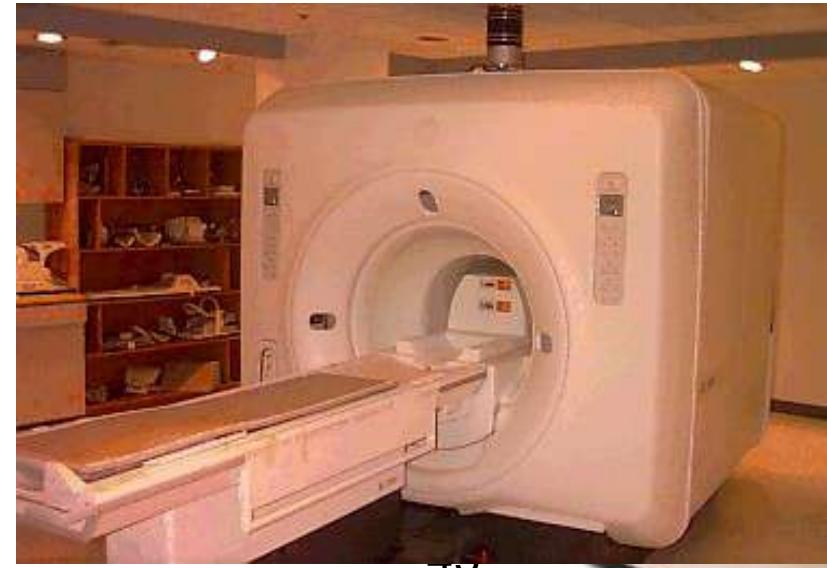
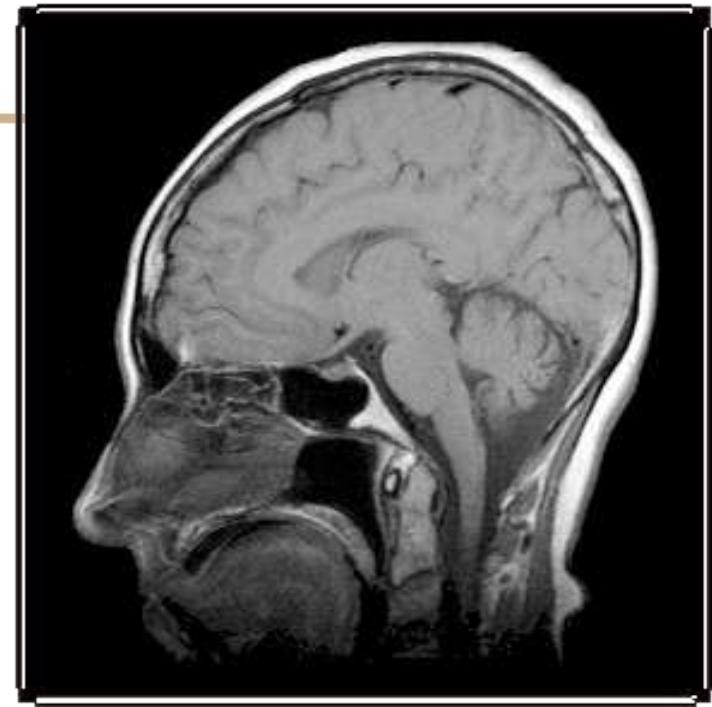
MRI 较 CT 具有独特的优点和特点：无电磁辐射损伤，对软组织具有更高的分辨率，多方向、多参数成像方法，无需用造影剂就能对心血管成像。





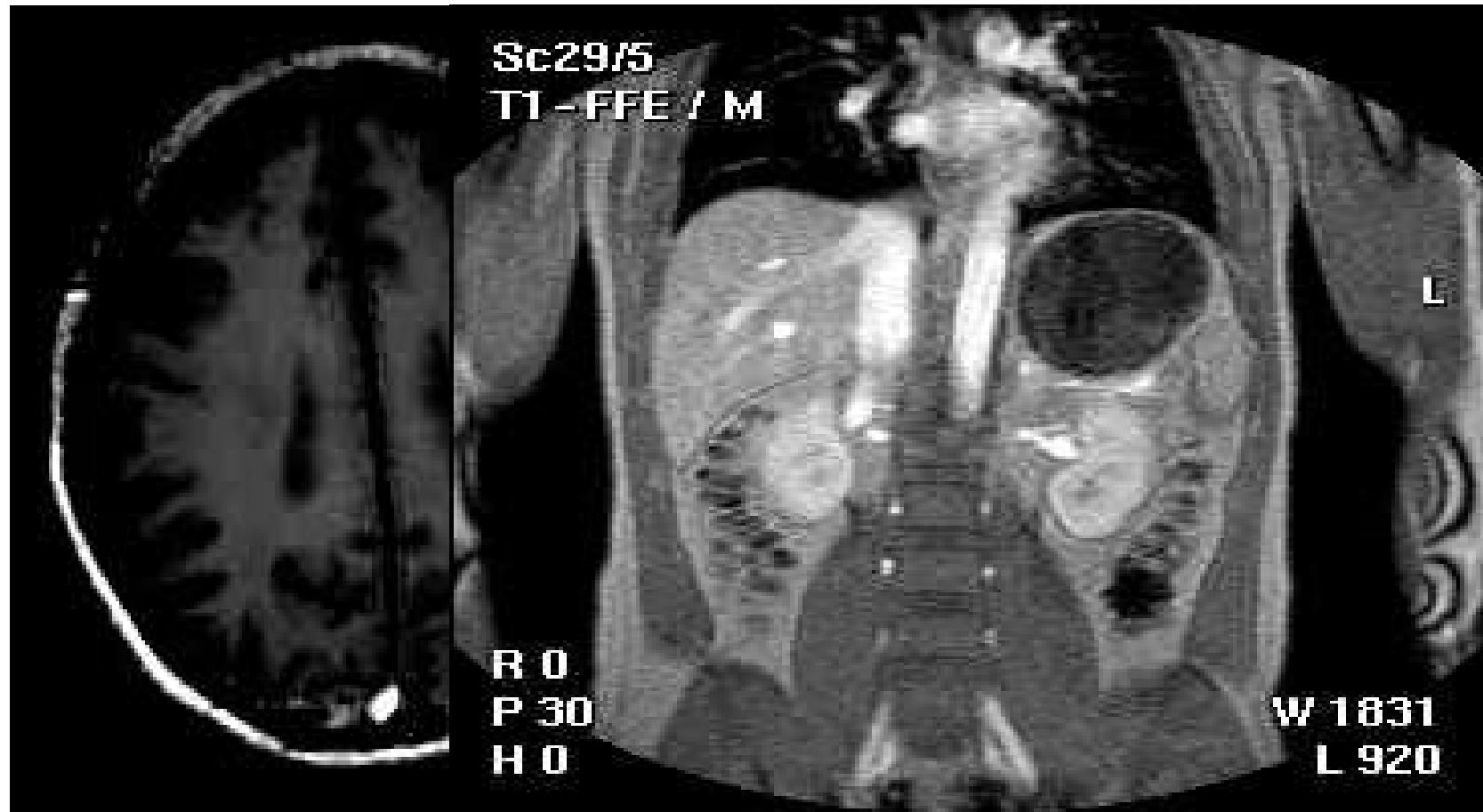


MRI





MRI 图像





fMRI 图像





医学图像的形式

- 根据用于成像的物质波的不同，医学图像分成五种形式：X线图像，放射性同位素图像，超声图像和磁共振图像。
- 根据成像设备是对组织结构（形态）成像还是对组织功能（代谢）成像，医学图像分成两类，即医学结构图像和医学功能图像
-

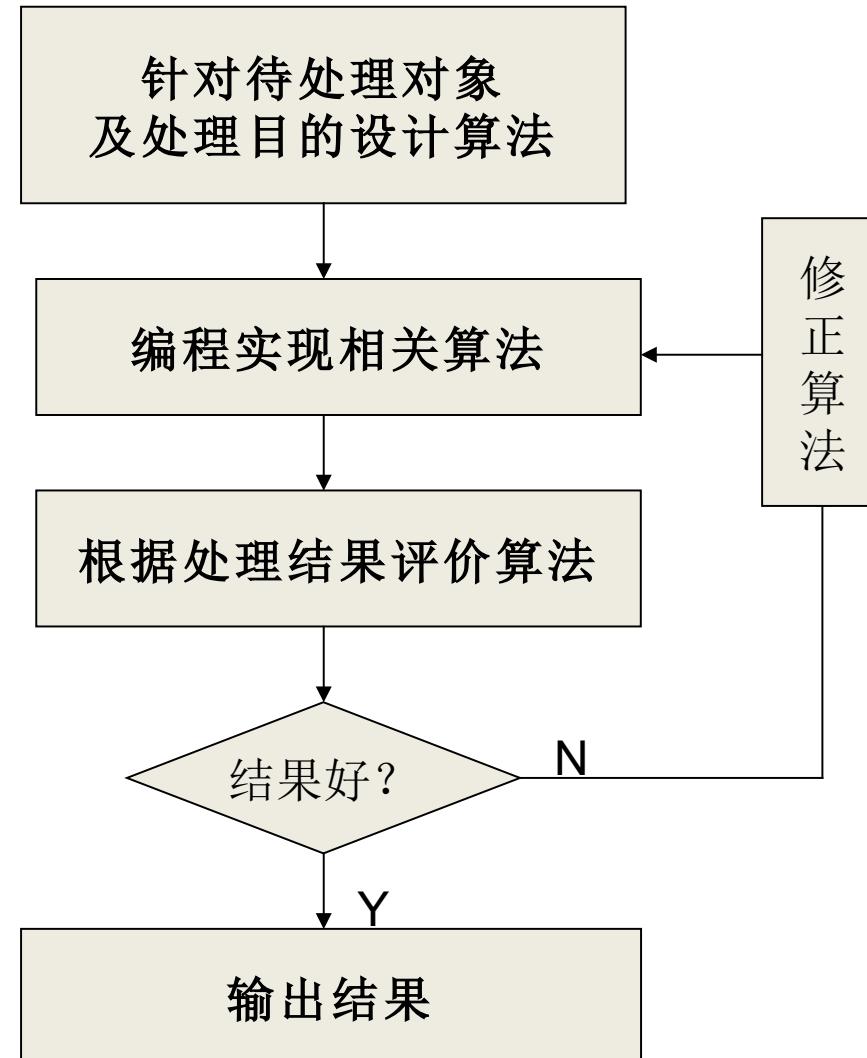


8. 5. 3 医学图像处理

医学图像的运算、医学图像变换技术、医学图像增强技术、医学图像分割技术、医学图像配准、医学图像三维可视化等。



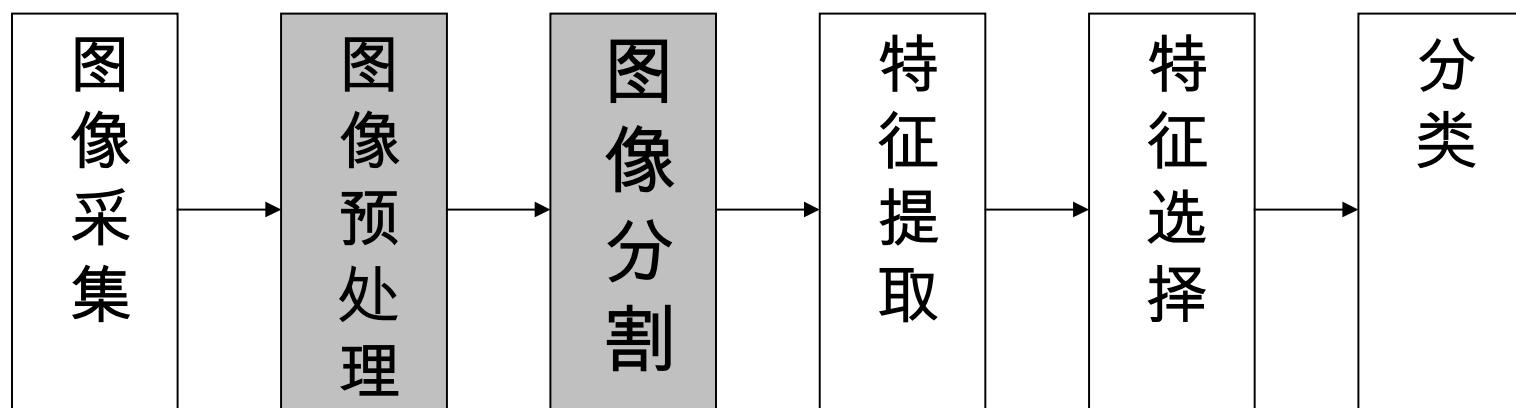
医学图像处理的基本步骤





8.5.3.1 乳腺肿瘤超声图像分割

乳腺癌诊断系统图示。





1. 预处理

1、要解决在滤除噪声的同时较好的保存图像的边缘、细节信息。超声图像去噪的目的是在平滑噪声的同时尽可能地保留原始图像的重要特征。比较常用的去噪方法如邻域平均法、低通滤波、维纳滤波等。

2、如何增强图像和改善图像的对比度及清晰度。超声图像在去噪过程中一些细节和边缘有或多或少的模糊，这就需对去噪后的图像进行增强。另外医学超声图像的对比度比较低，图像往往偏暗或偏亮，这就需改善图像对比度

ROI 区域选取

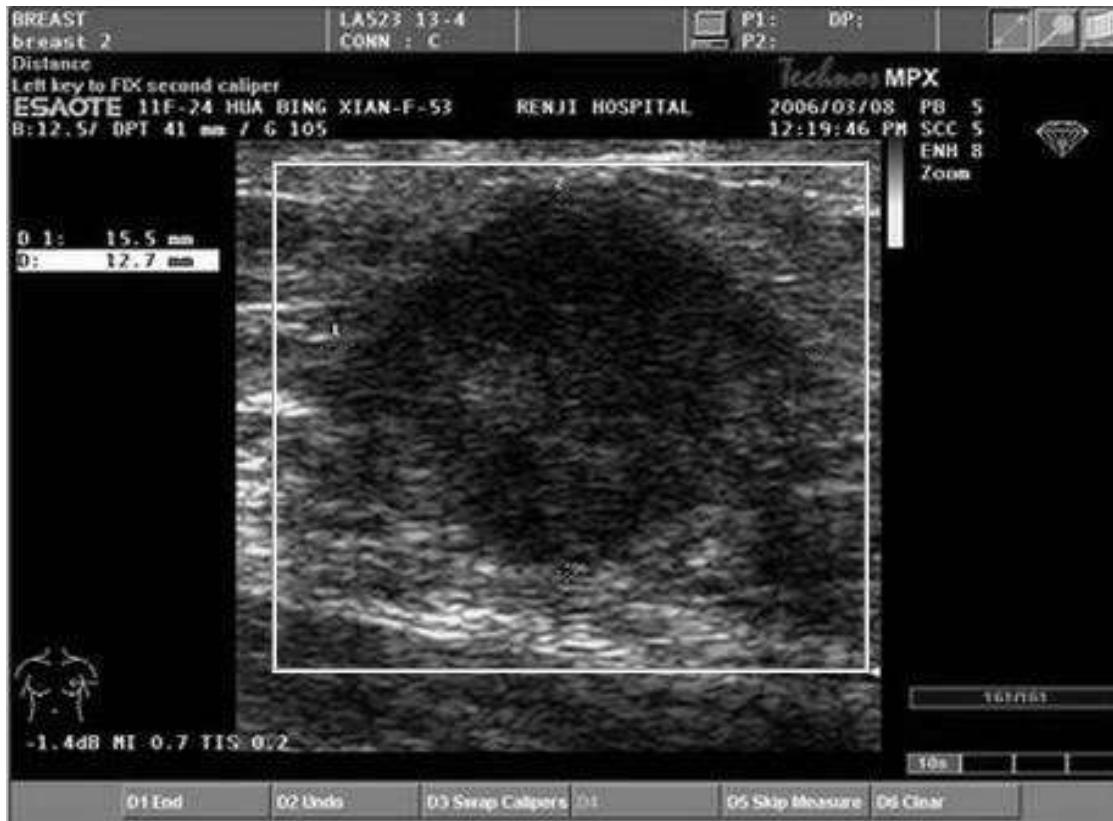
去噪处理

图像增强

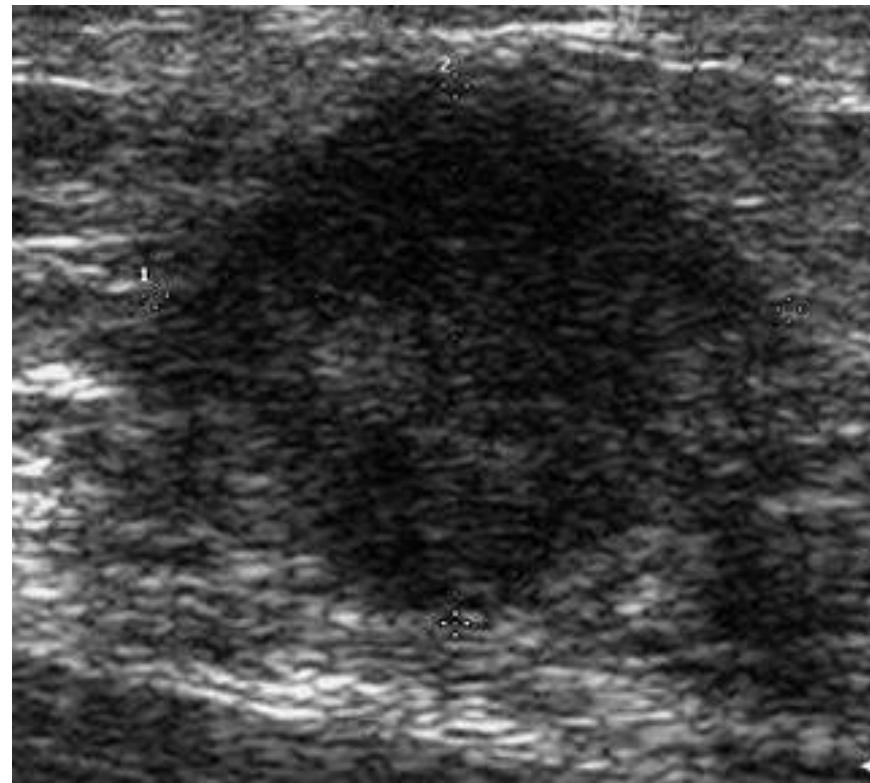


ROI 区域选取

感兴趣区域为包含肿瘤的尽量小的区域



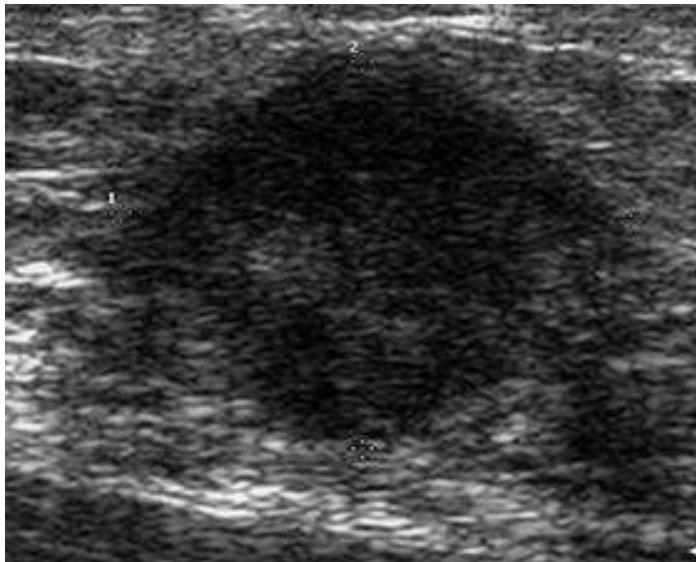
原始乳腺肿瘤超声图像



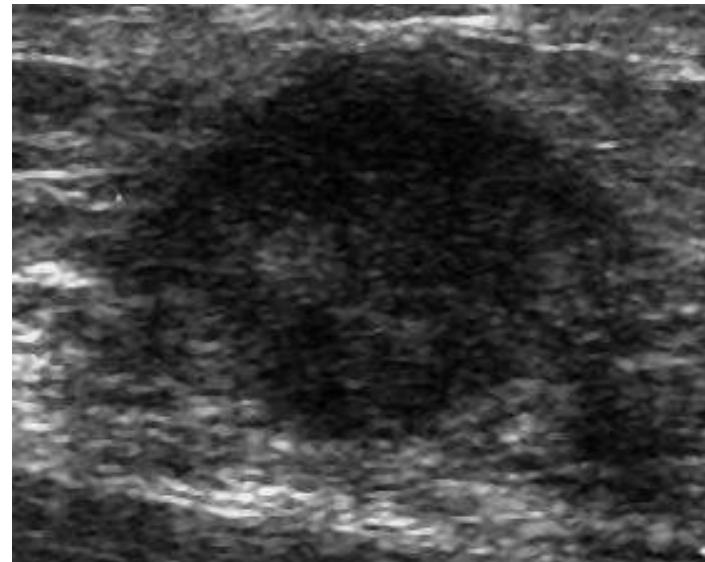
ROI 区域图像



使用中值滤波算法进行去噪处理



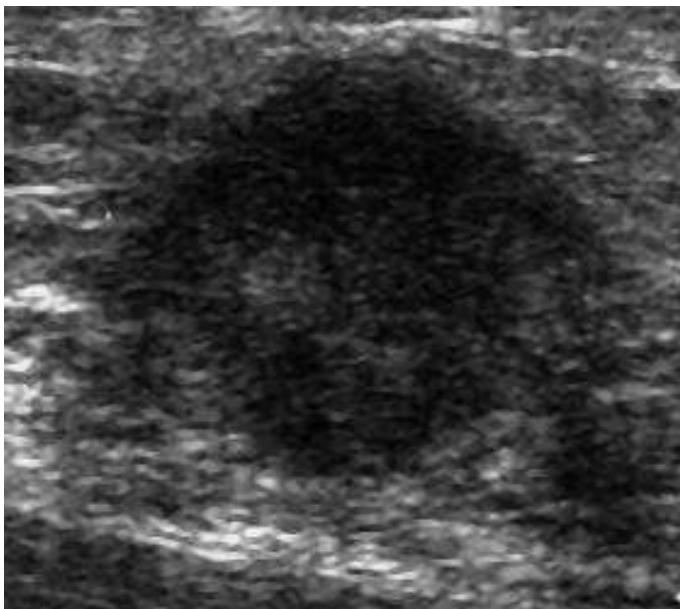
原始图像



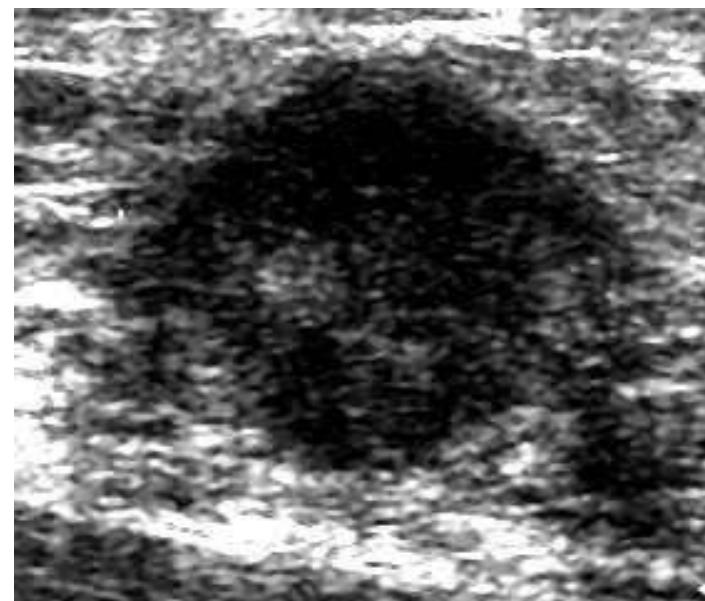
中值滤波图像



对比度增强效果



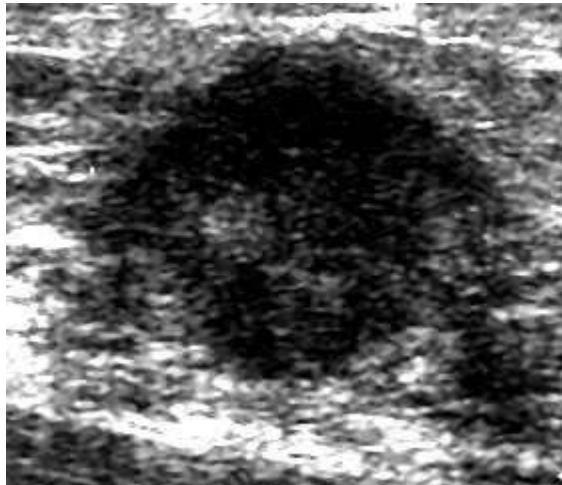
对比度增强前图像



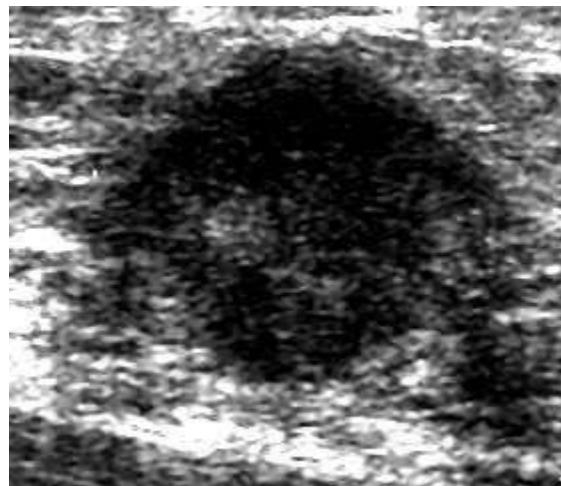
对比度增强后图像



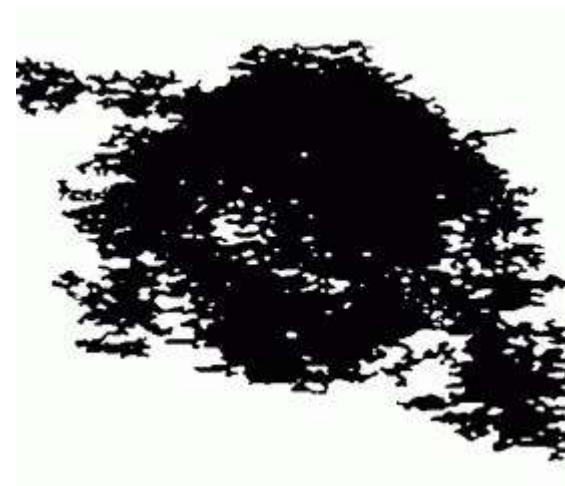
2. 基于乳腺超声图像的肿瘤区域分割



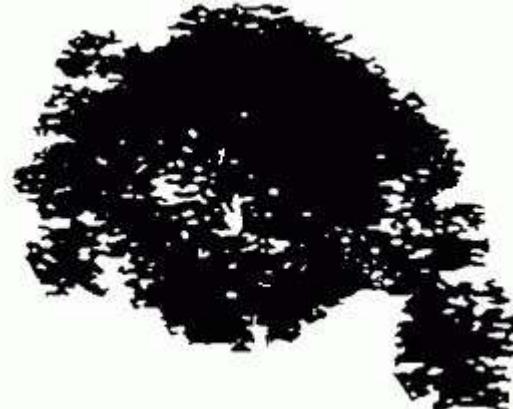
二值化后图像



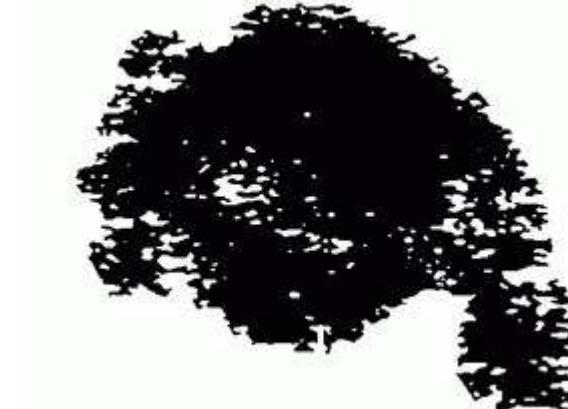
(a) 预处理后图像



(b) 基于已生长区域阈值的生长结果



(c) 基于局部统计值阈值的生长结果



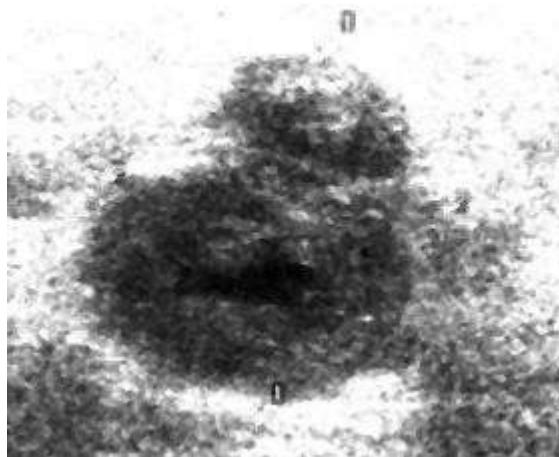
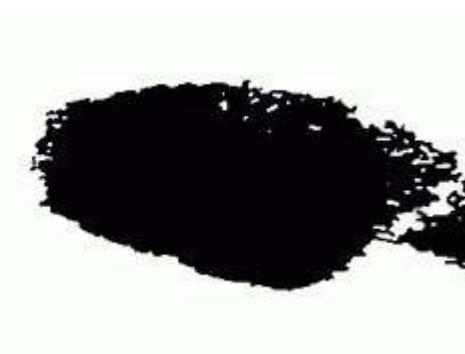
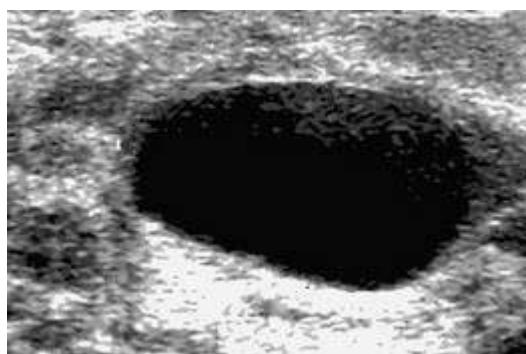
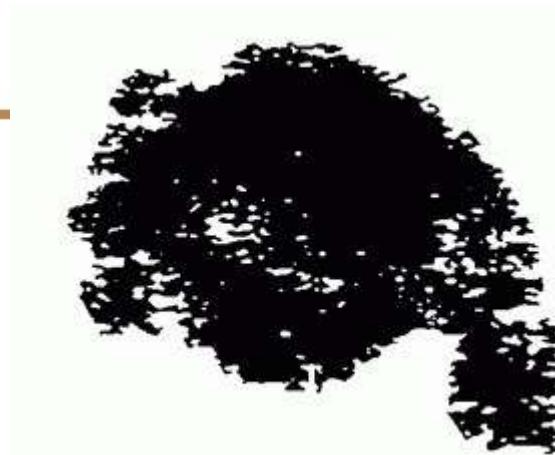
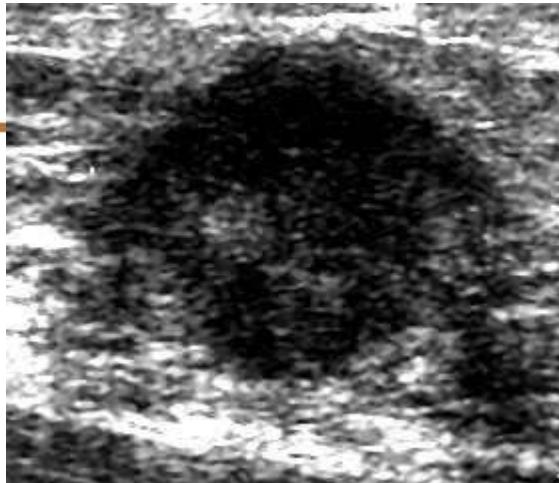
(d) 基于已生长区域特征的阈值选取生长结果

基于已生长区域的生长结果内部比较均匀，但它的边界上较多毛刺，对下一步边界提取干扰很大；基于局部统计值的生长区域边界较平滑、毛刺少，接近实际肿瘤边界，但内部空洞较多；综合二者特征的生长区域不仅内部较好，且边界毛刺少，较平滑，可以较为准确的提取出肿瘤边界。



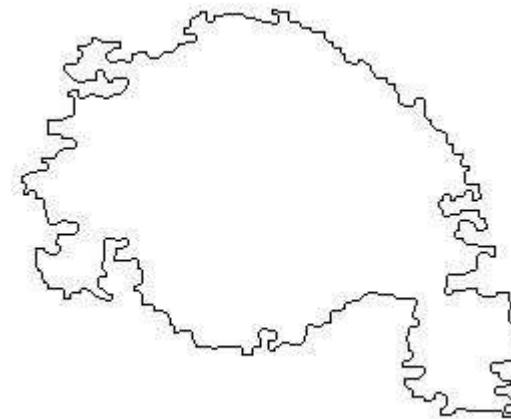
- 基于区域生长的乳腺肿瘤初始区域获取方法。

先采用阈值化的方法粗略的对图像进行分割，然后以二值化后图像的黑色像素的重心点为中心，选取一个大小的窗口，计算该窗口内所有像素灰度值的平均值，选择窗口内灰度值最接近平均值的像素点作为区域生长的种子点。以种子点附近的统计值和已生长区域特征相结合的方法确定生长阈值，综合二者特征的生长区域不仅内部较好，且边界毛刺少，较平滑，可以较为准确的提取出肿瘤边界。





3 基于形态学与边界跟踪的肿瘤边缘提取

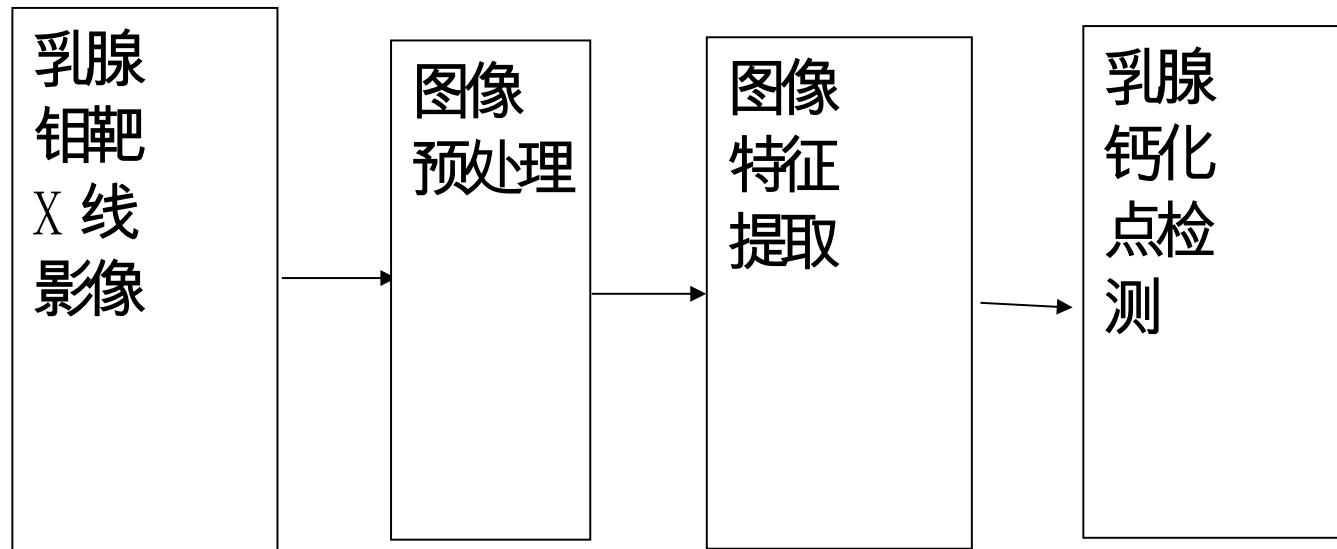


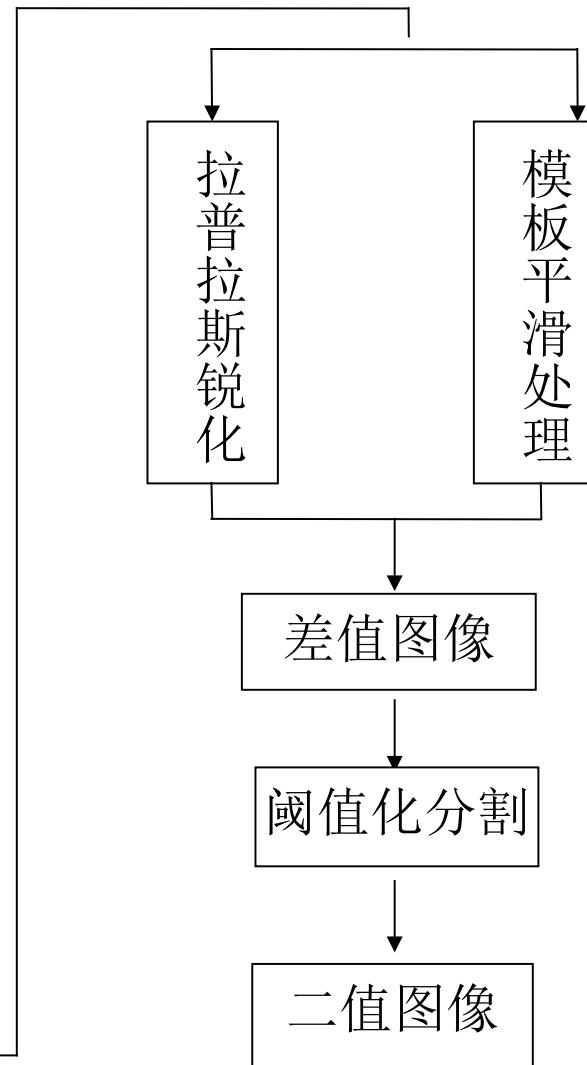
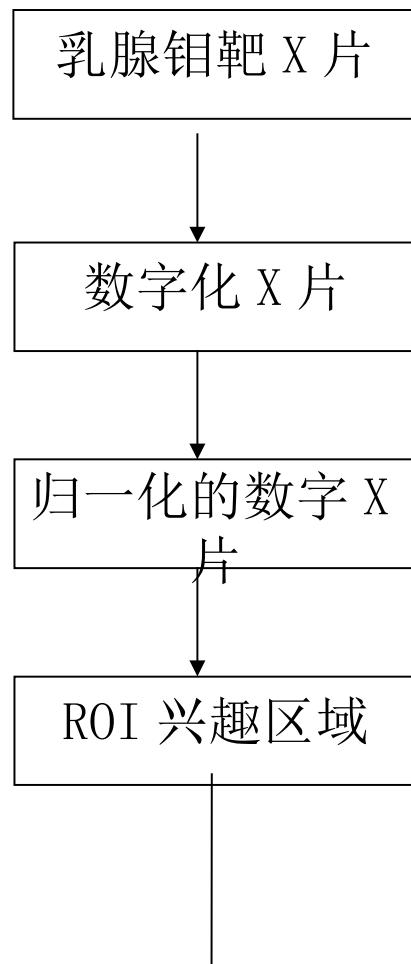


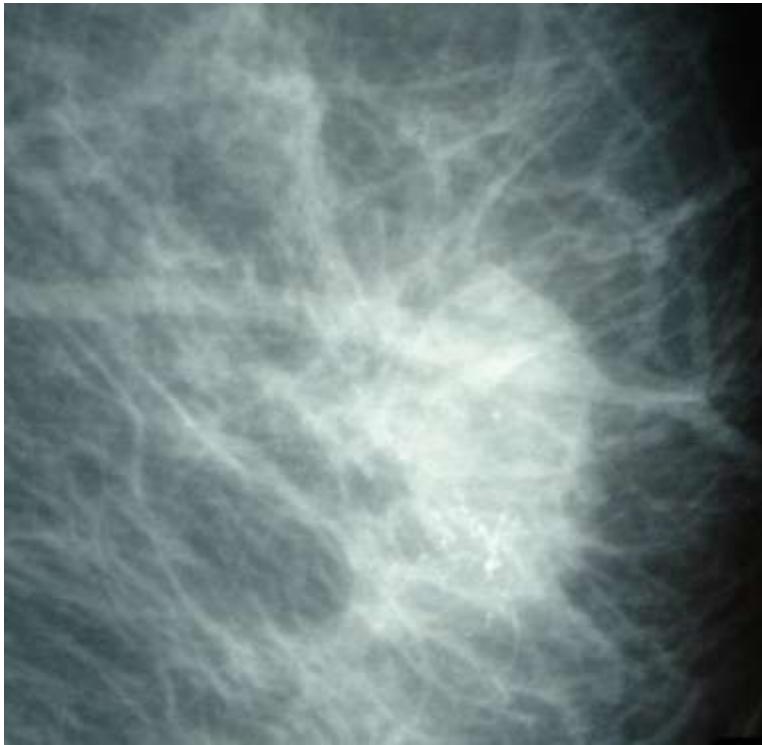
- 运用形态学方法对初始肿瘤区域进行滤波，采用了使用 3×3 模板先开后闭 形态学滤波运算，使肿瘤边缘的毛刺得到消除，肿瘤中间的空洞也得到了填充，同时图像中粘连的非肿瘤区域也被分割出来。
- 改进的边界跟踪算法，算法中给出了位置编码的定义，根据当前点在上一边界点的位置，确定搜索下一目标点时的起始点，因此在搜索下一目标点时不必对 8 邻域的点全部搜索，最多只需要搜索 5 次便可以找到目标点。这种方法在准确跟踪到肿瘤边缘的同时，缩短了边界跟踪的时间



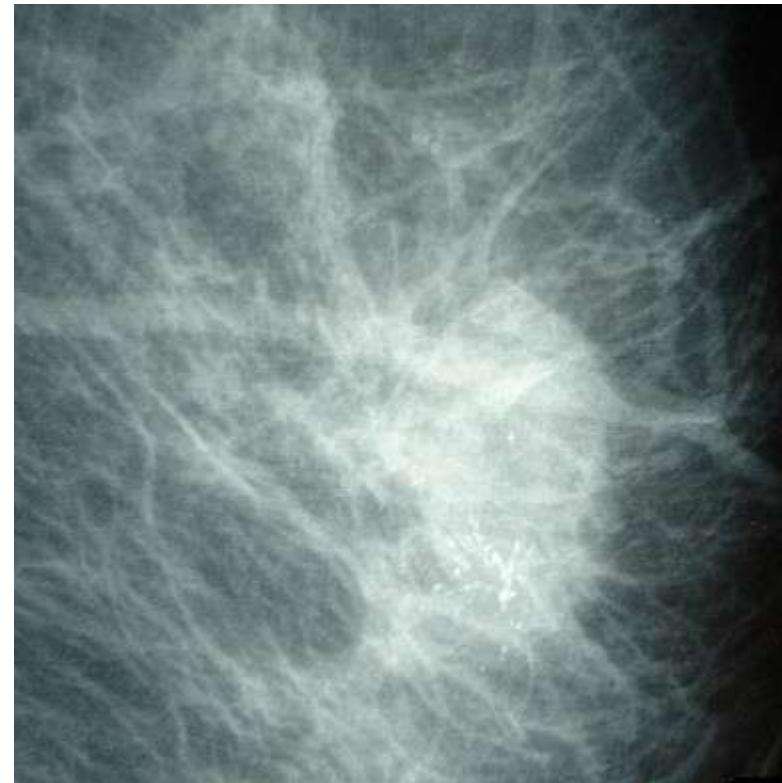
8. 5. 3. 2 乳腺微钙化点自动检测方法



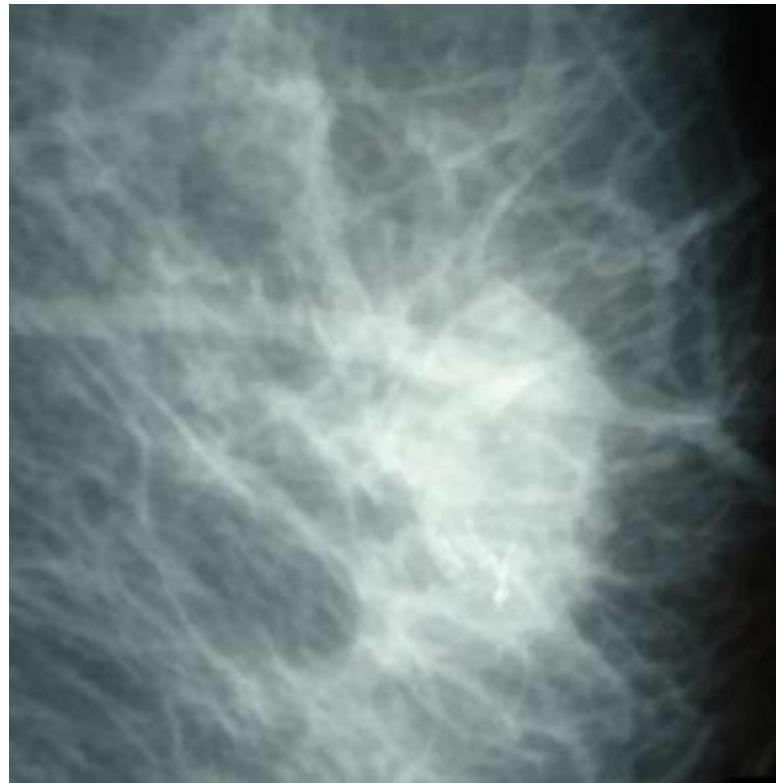




原始图像



拉普拉斯增强效果图



平滑滤波效果图



差值图像



分割图像

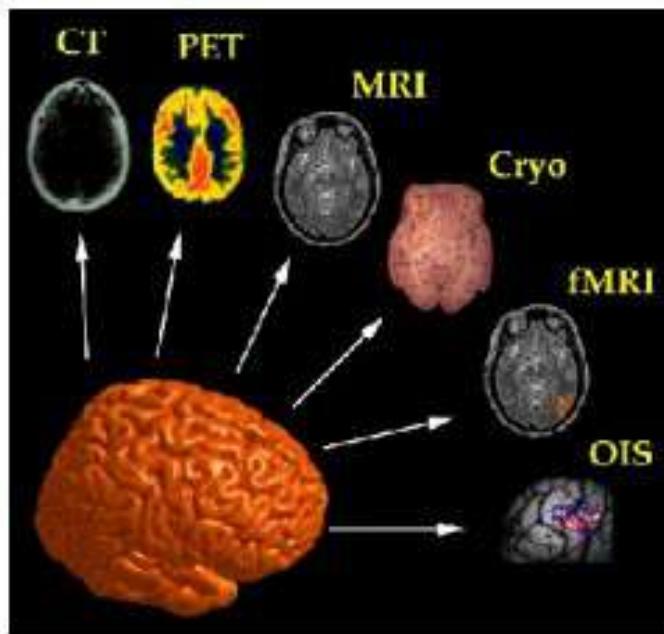


修正后的分割图像



8.5.3.3 医学图像配准

由于成像的原理和设备不同，存在有多种成像模式（解剖成像、功能成像）。



对人脑可用多种模式成像

配准的目的：将多种成像模式或同一种模式得到的多幅图像综合分析。

解决的问题：几幅图像的严格对齐



配准的概念

- 对几幅不同的图像作定量分析，首先要解决这几幅图像的严格对齐问题，这就是我们所说的图像的配准（image registration）。
- 医学图像配准是指对于一幅医学图像寻求一种（或一系列）空间变换，使它与另一幅医学图像上的对应点达到空间上的一致。这种一致是指人体上的同一解剖点在两张匹配图像上有相同的空间位置（位置一致，角度一致、大小一致）。
- 几幅图像信息综合的结果称作图像的融合（image fusion）。



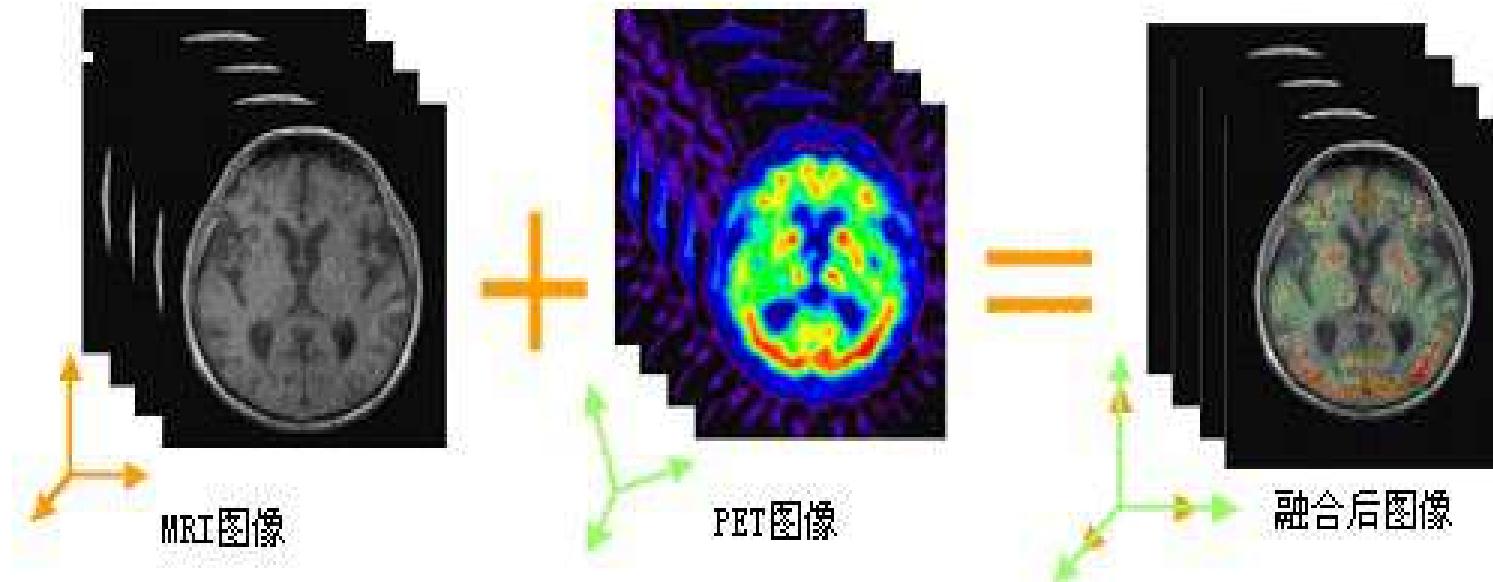
图像配准示意图（从不同角度，不同位置得拍摄的照片，每张照片反映某些方面的特征）：

配准实现过程：

1. 获得特征图像：不同角度、不同条件
2. 图像配准：通过空间变换（移动和旋转），使两幅图像对齐
3. 图像融合：得到整体特征图像



实例：



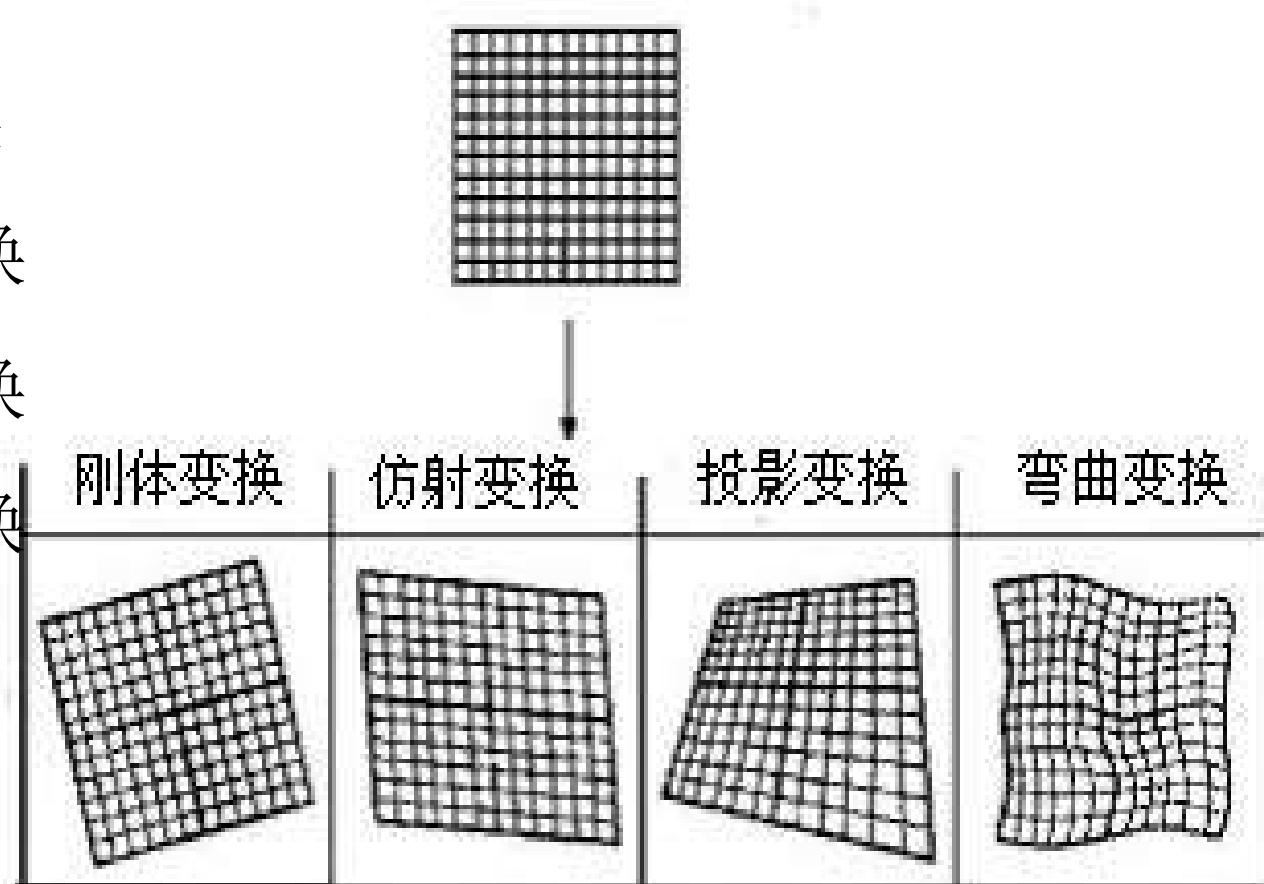
8.5.3.3 节图片内容选自以下网站教学资源

<http://cc.usst.edu.cn/able.acc2.web/template/view.aspx?wmz=4074&cou rsetype=0&courseid=275&topmenuid=48976&menutype=4&contentid=4 074&action=view&type=&name=&menuid=000002>



几何变换的种类

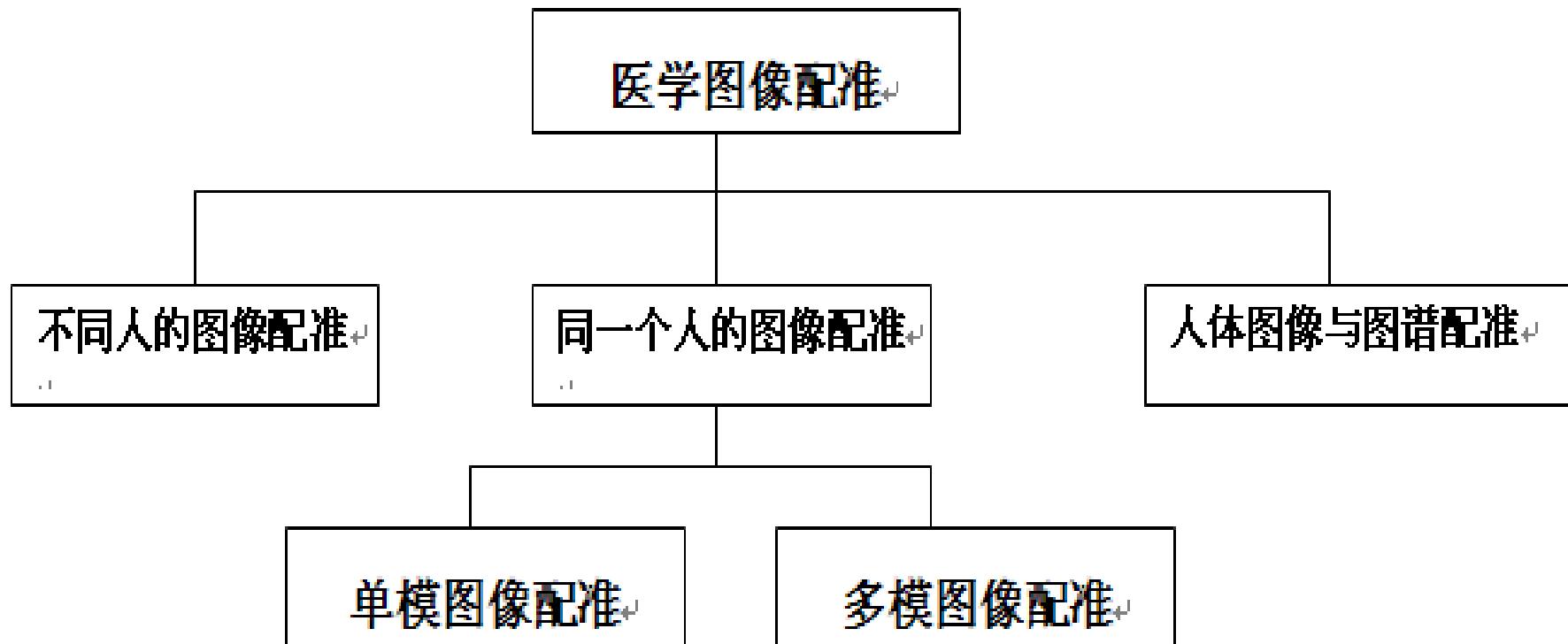
- 刚体变换
- 仿射变换
- 投影变换
- 弯曲变换





配准的类型

根据成像模式的不同，以及配准对象间的关系等，医学图像配准可有多种不同的分类方法：





1. 同一对象的图像配准

- 单模图像配准：相同的成像设备，不同角度或时间或不同成像条件
- 多模图像配准：不同的成像设备



单模配准的典型应用：

(1) 不同 MR 加权像间的配准

T1 加权像
T2 加权像
质子密度加权像 } 不同组织表现不同强度 → 信息互补

(2) 电镜图像序列的配准

不同时间采集的多幅图像 → 粒子移动，形态变化
(时间序列图像)
→ 研究生长现象



单模配准的典型应用：

(3) fMRI 图像序列的配准

时间序列图像，大脑活动会产生氧摄取量和血流间的不平衡

————→ MRI 图像测得信号的改变

(4) 胸腹部脏器的图像配准

不同时刻的三维脏器图像 ——→ 四维图像分析

解决：参考点位置确定（定位问题）

运动过程中的局部变形和噪声问题



多模医学图像配准：

待配准的两幅图像来源于不同的成像设备。

解剖图像：CT：对密度差异较大的组织效果好
MRI：可识别软组织

较高的空间分辨率

功能图像：SPECT/PET：能反映人体的功能和代谢信息

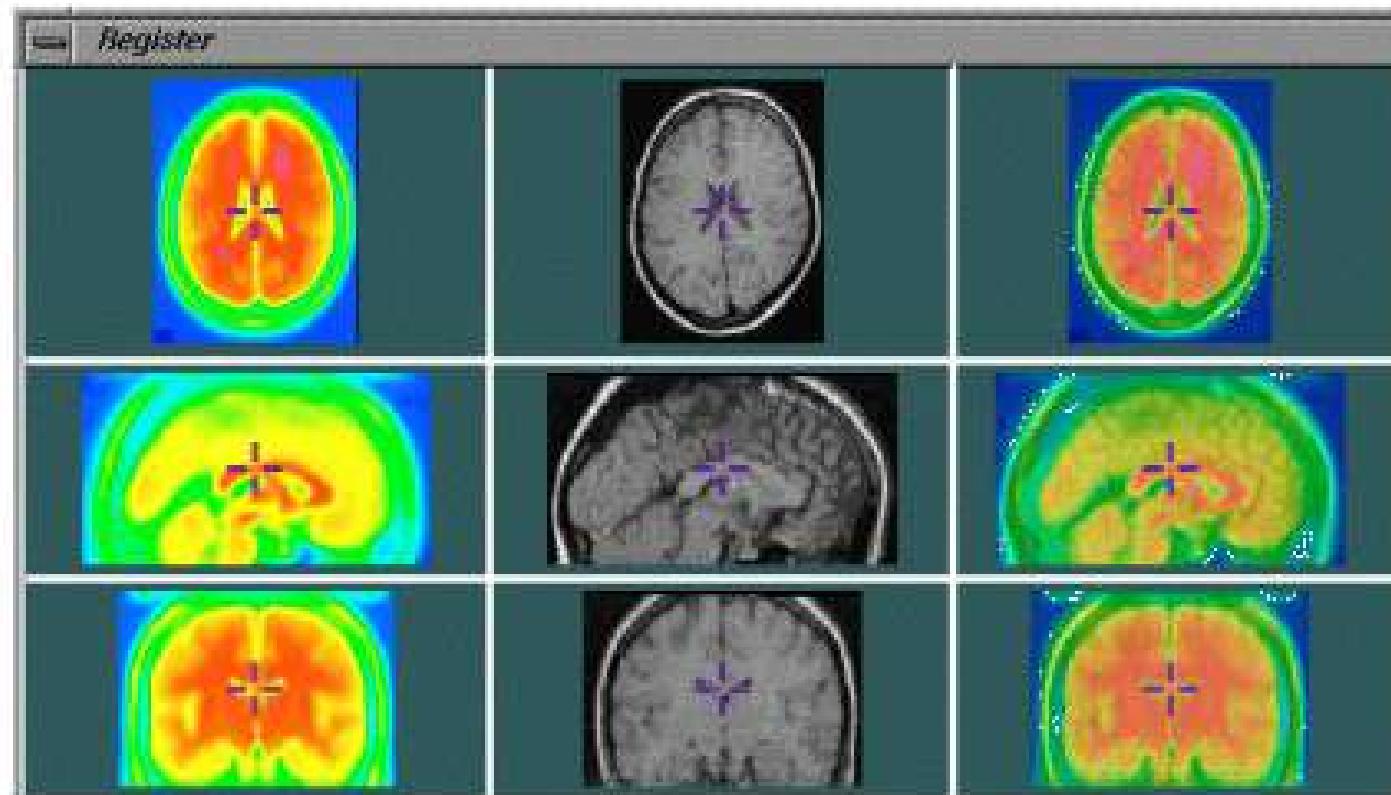
空间分辨率差

一般临床上的应用：

CT（或MRI）和SPECT（或PET）的综合分析，同时提供功能和解剖信息



多模医学图像配准：



人脑 MR/PET 图像配准

上排：轴向图； 中间：矢状图； 下排：冠状图
左： PET 图像； 中： MRI 图像； 右：融合图像

难点：由于扫描设备的原理不同，两种断层图像间并不存在简单的一一对应关系



8.5.3.4 基于 CT 数据的三维重建与切片剪取

- VTK 可视化类库
- 体绘制技术的基础知识、实现过程和主要算法
- 光线投射算法的实现过程
- 重建实体的图像分割和切片剪取技术



VTK 可视化类库

- VTK (The Visualization Toolkit) 是一个开源的、可以免费获取的、关于 3D 计算机图形学的、图像处理、计算机可视化的类库。
- VTK 可视化模块包含数据对象 (data objects) 和处理对象 (process objects)。
- 处理对象被分为源 (source) 、过滤器 (filter) 和映射器 (mapper)。



三维重建和体绘制技术

- 三维重建是针对一系列二维图像，通过重建算法，重建出三维实体，并将其直观的显示出来，供用户直接查看和交互操作的可视化方法。
- 体绘制技术也被称作直接体绘制方法，采用直接对所有的二维数据进行明暗处理的方法，最后通过合成得到具有三维效果的图像。



体绘制技术的特点和过程

- 体绘制的特点：
 - 1、调整颜色和透明度，显示和区别不同的部位。
 - 2、绘制物体的表面和内部细节信息。
 - 3、不借助中间几何图元，直接将三维数据场绘制到二维图像屏幕上。
- 体绘制的过程：
 数据获取、数据处理、坐标变换、体素绘制和累加和显示

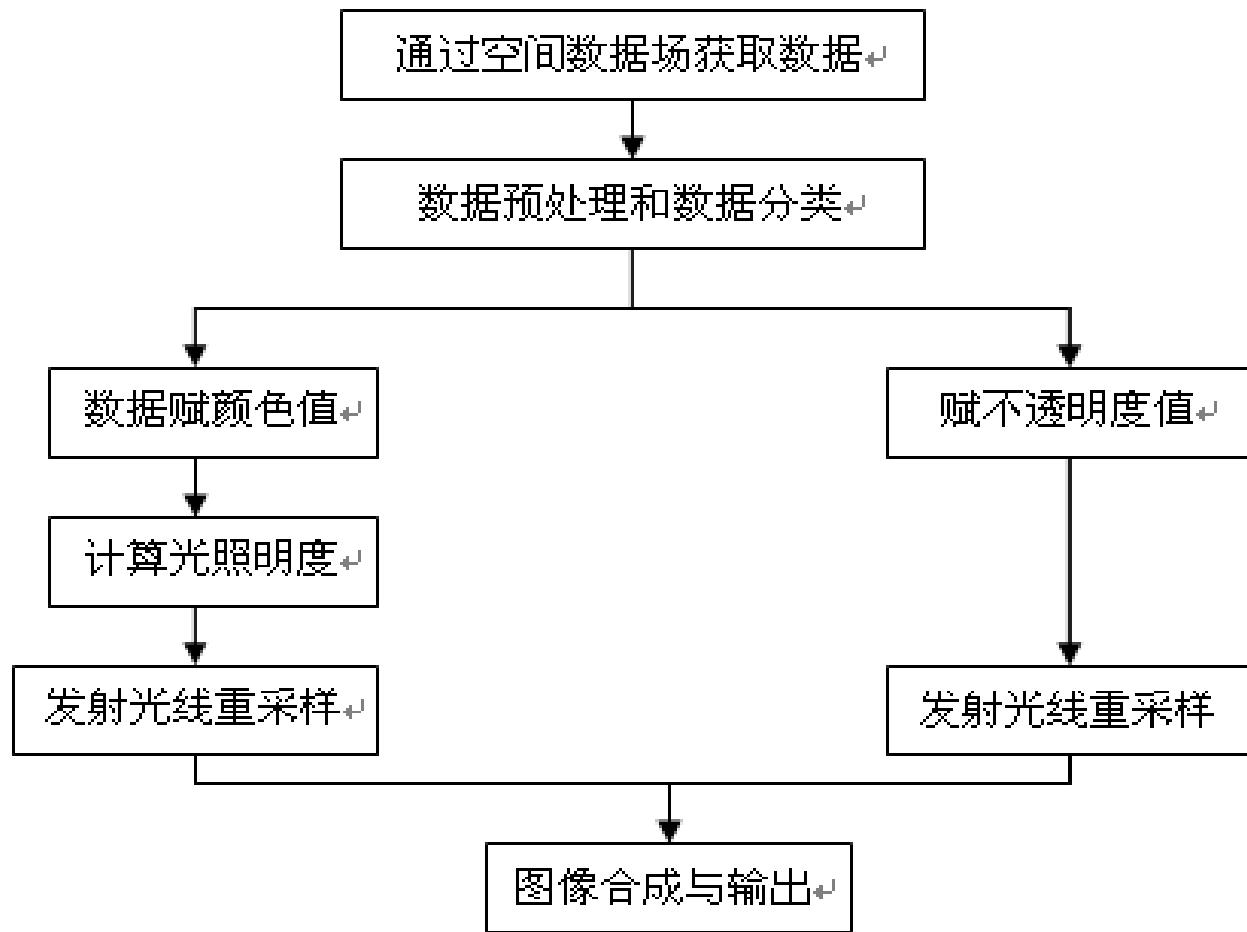


体绘制的算法 ---- 光线投射算法的原理

- 对图像的每个像素，沿固定的视线方向发射一条光线；
- 光线穿过三维数据场，对数据场进行采样以获取颜色和透明度信息；
- 根据光线吸收模型对颜色值和透明度进行累加，直至光线从数据场中穿出；
- 合成得到最后的三维图像。



光线投射算法流程图





数据分类

- 本文数据场由 460 幅头部 CT 图像序列构成；
- 三维数据场用集合 D 表示，数据分类如下：

$$\begin{cases} D = \bigcup_{i=0}^n D_i \\ D_i \cap D_j = \emptyset (0 \leq i < j \leq n) \end{cases}$$

- 目前常用的分类方法有阈值法和概率法。



数据赋颜色值和透明度值

- 数据分类结束后，数据场中的 CT 图像仍是灰度图像。为了增强细节显示效果和保证信息量，对图像进行伪彩色处理，根据线性映射函数，将灰度图像映射为彩色图像。
- 对透明度采用灰度值线性化方法，使透明度映射为 0-1 之间的连续数值。值为 0 时，表示完全透明；值为 1 时，表示完全不透明。



射线穿越体数据

- 为数据设置好颜色值和透明度值后，需要沿射线的方向穿越体数据并完成采样。
- 沿射线方向穿越体数据：
 - 1、 穿越体数据的过程可以看成光线穿越确定了顶点纹理坐标的三维立方体的过程；
 - 2、 在整个穿越过程中，计算采样体数据坐标，并进行体数据采样；
 - 3、 当光线投出立方体或者累加的透明度为 1 时，采样过程结束。



三维数据场的重采样过程

- 使用离散的三维数据场构造连续的三维数据场；
- 根据采样点的分辨率，计算出被采样信号的奈魁斯特 (Nyquist) 频率极限，采用低通滤波函数去掉高于这一极限频率的成分；
- 对滤波后的函数进行重新采样。



图像合成

- 图像合成：沿着从像素点发出的射线，计算出的该射线上各采样点的颜色和不透明度，按照一定的合成规则，形成像素点的最终颜色值和不透明度值，完成图像合成。
- 1、由前向后合成：

$$C_i^{\Delta} = (1 - A_{i-1}^{\Delta}) C_i + C_{i-1}^{\Delta}$$

$$A_i^{\Delta} = (1 - A_{i-1}^{\Delta}) A_i + A_{i-1}^{\Delta}$$

- 2、由后向前合成：

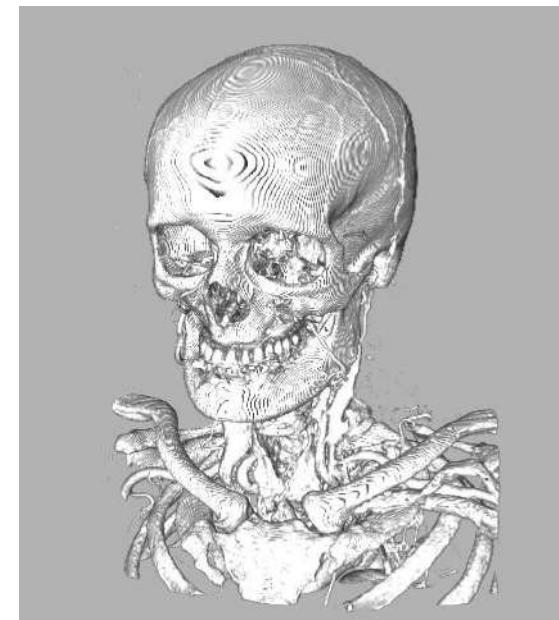
$$C_i^{\Delta} = (1 - A_i) C_{i+1}^{\Delta} + C_i$$

$$A_i^{\Delta} = (1 - A_i) A_{i+1}^{\Delta} + A_i$$



绘制结果

本文采用 460 帧 $512 \times 250 \times 12\text{bit}$
的 CT 头部图像
序列作为体绘制的实验数据，重建效果如下：





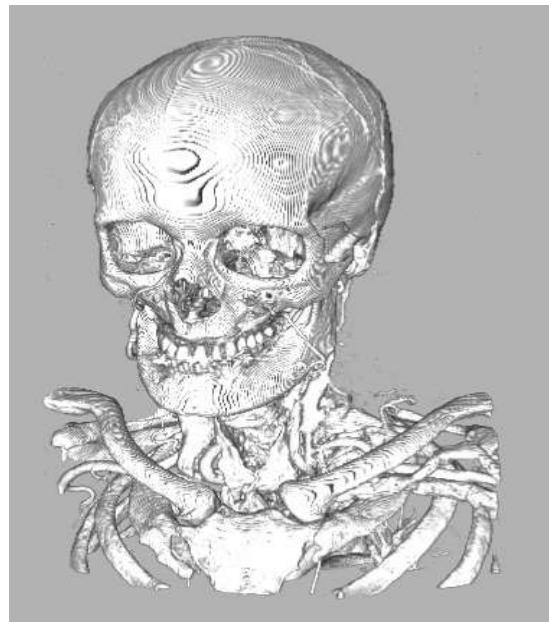
光线投射算法的改进

- 为了满足三维重建的实时性要求，对光线投射算法进行改进，加快重建的速度，提高重建的实时性。
- 光线投射算法改进的方法：
 - 1、包围盒裁剪数据场
 - 2、投射光线离散化
 - 3、提前不透明度截止



重绘制

采用立方体包围盒裁剪数据场方法，对三维数据场进行裁剪，并调整了抽样距离，对算法进行了改进。采用相同配置的 PC 进行测试，改进后的三维重建耗时为 51.20 秒，速度提高了 8.35 秒，使重建的效率提高了 13%，并且重建的效果没有发生明显的变化。改进后效果如图所示：





重建实体的分割剪取技术

- 为了更好的显示内部结构的细节信息，研究了重建实体的分割剪取技术。
- 图像分割技术通过构造分割子体，得到感兴趣部位的体细节信息。
- 切片剪取技术通过构造剪取平面，得到感兴趣部位的面细节信息。



头部实体分割

- 由于人脑分为左半球和右半球，在观察头部实体时，分为左部和右部进行重建和观察。
- 效果如图所示：





切片剪取的过程

- 1、三维重建获得三维体数据；
- 2、需要确定剪取平面，也就是切片所在的平面。直接确定剪取平面是不容易的，一般通过确定剪取平面的法向量和平面与平面上的一个点，通过点法式来确定剪取平面；
- 3、通过剪取函数载入选定的剪取平面，完成三维体数据的切片剪取过程。



VTK 中完成切片剪取

- 1、实例化一个 `vtkPlane` 对象，增加一个剪取平面，同时为此平面定义相关参数（主要包括法向量和平面上的一点）；
- 2、使用 `AddClippingPlane()` 方法将剪取平面载入 `mapper` 对象中。在这个过程中，添加了标识两个平面的 `vtkPlane` 对象，这两个平面是平行的，并且中间距离非常微小，就构成了一个厚度非常薄的长方体。
- 因此，切片剪取的过程可以看作是将这个长方体中的三维体数据显示出来的过程。



冠状面上切片剪取

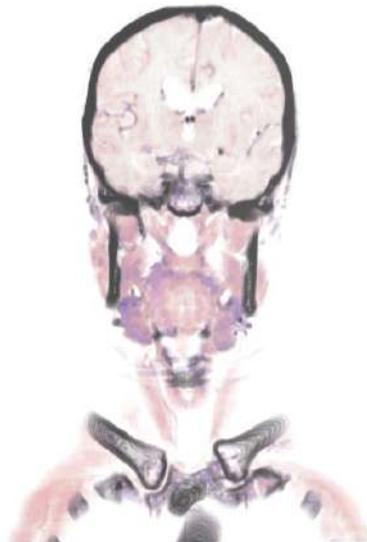
- 冠状面的法向量都是 $(1,0,0)$ 和 $(-1,0,0)$ ，冠状面上的点取 x 轴上的点，坐标为 $(,0,0)$, 其中的取值范围是 0-200 , 当取 遍 0-200 的所有值时，通过构造两个间距非常小（程序中
间距为 6 ）“ ” 的平行平面，可以得到所 有的冠状面





矢状面上切片剪取

- 矢状面的法向量都是 $(0,1,0)$ 和 $(0,-1,0)$ ，矢状面上的点取 y 轴上的点，坐标为 $(0, ,0)$, 其中的取值范围是 0-200，当取遍 0-200 的所有值时，通过构造两个间距非常小（程序中间距为 6）的法向量成 180 度角的平行平面，可以得到所有矢状面」





任意平面上切片剪取

- 法向量的确定

$$\cos \alpha = \alpha_1 / |\text{OP}| = \alpha_1 / \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}$$

$$\cos \beta = \alpha_2 / |\text{OP}| = \alpha_2 / \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}$$

$$\cos \gamma = \alpha_3 / |\text{OP}| = \alpha_3 / \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}$$

- 切面内点的选择

为了方便计算，我们可以在确定切片平面的法向量之后，

沿坐标轴选取切面内点。当内点的坐标取遍坐标轴上所有

的存在相交切片的值时，就得到了这个法向量方向上的所

有切片。因此，实现实面上的切片剪取，只需要求出切片

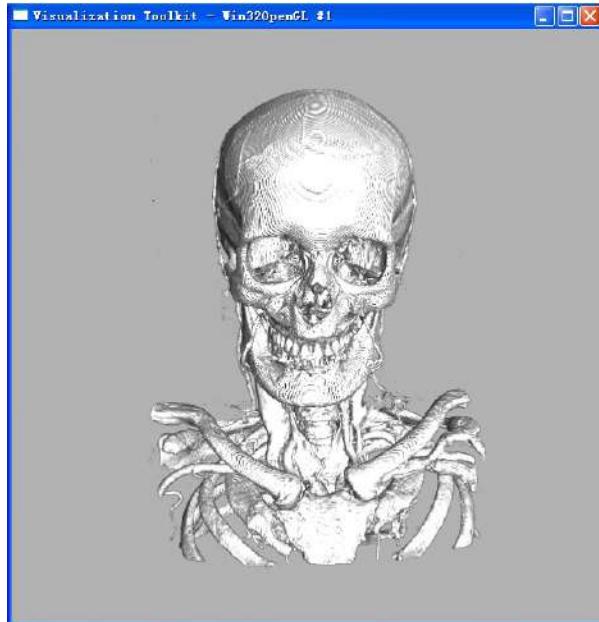


三维重建与分割剪取系统

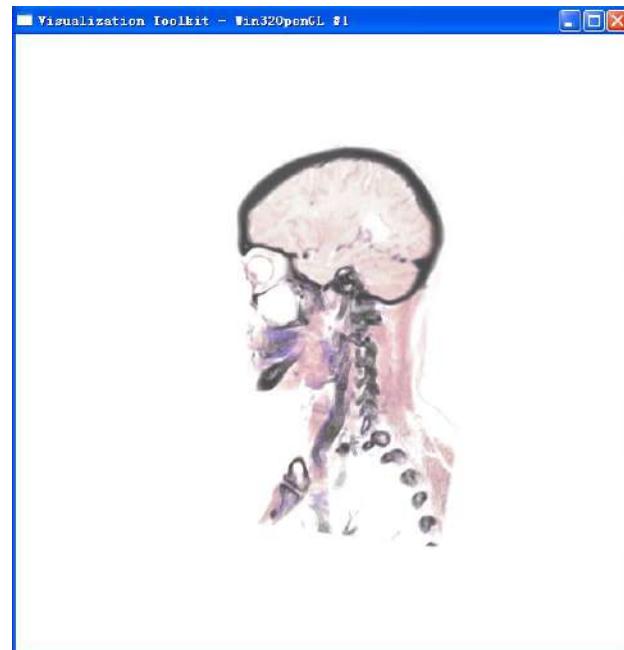
- 采用 VTK 可视化工具，采用 C++ 开发语言，利用 Microsoft Visual C++ 6.0 进行系统开发，作为 PACS 系统的三维重建模块，实现了 CT 图像的三维重建、显示、图像分割和切片剪取等功能。
- 系统具备下列功能：
 - 1、支持 DICOM 格式，可以读取指定路径下的 DICOM 数据集；
 - 2、根据光线投射算法，实现 CT 图像的体绘制三维重建；
 - 3、支持对重建后的三维实体进行图像分割操作；
 - 4、支持对重建后的三维实体进行切片剪取操作。



系统界面



冠状面切片剪取效果





数字图像处理

(Digital Image Processing)

山东科技大学
王萍教授



第8章 - 第3讲 遥感图像应用系统

- **遥感图像获取**
- 遥感、遥感信息、遥感的技术系统
- 遥感数据的获取方式及数据特点
- **遥感图像处理**



遥感图像获取

1 遥感概念： Remote Sensing

美国学者普鲁伊特于 1960 年提出， 1961 年 “环境遥感讨论会” 正式通过。

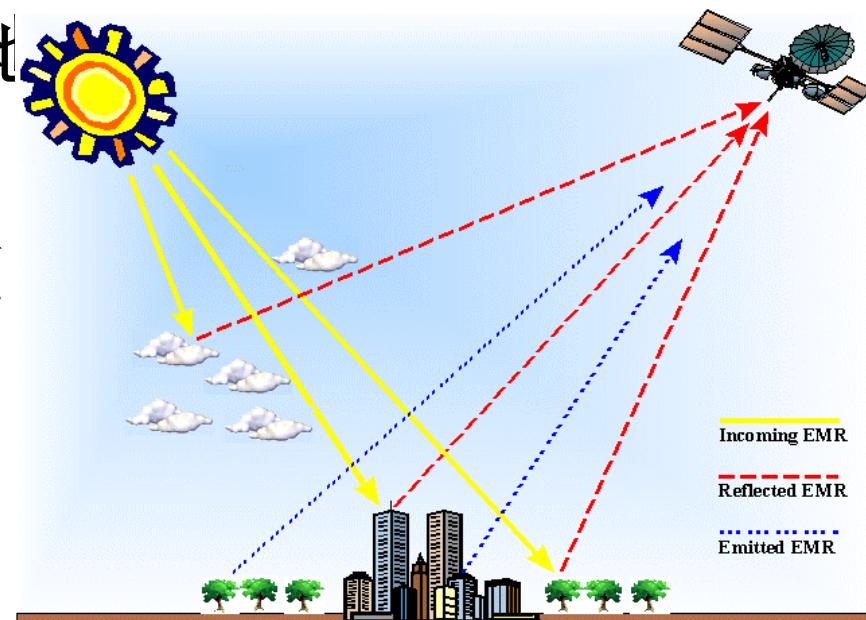
- 广义遥感：
 - 泛指一切无接触的远距离探测，包括对电磁场、力场、机械波（声波、地震波）等的探测。
- 狹义遥感：
 - 是指从远离地面的不同工作平台上（气球、飞机、人造卫星、航天飞机），通过传感器对地球表面的电磁波信息进行探测，并经信息的传输、处理和判读分析，对地球的资源和环境进行探测的综合性技术。



遥感图像获取

2 遥感信息 (RS Information)

- 利用安装在遥感平台上的各种电子和光学传感器，在高空或远距离处
- 接收到的，来自地
- 面或地面以下一定
- 深度的地物反射或
- 发射的电磁波信息



来源互联网：不详



遥感图像获取

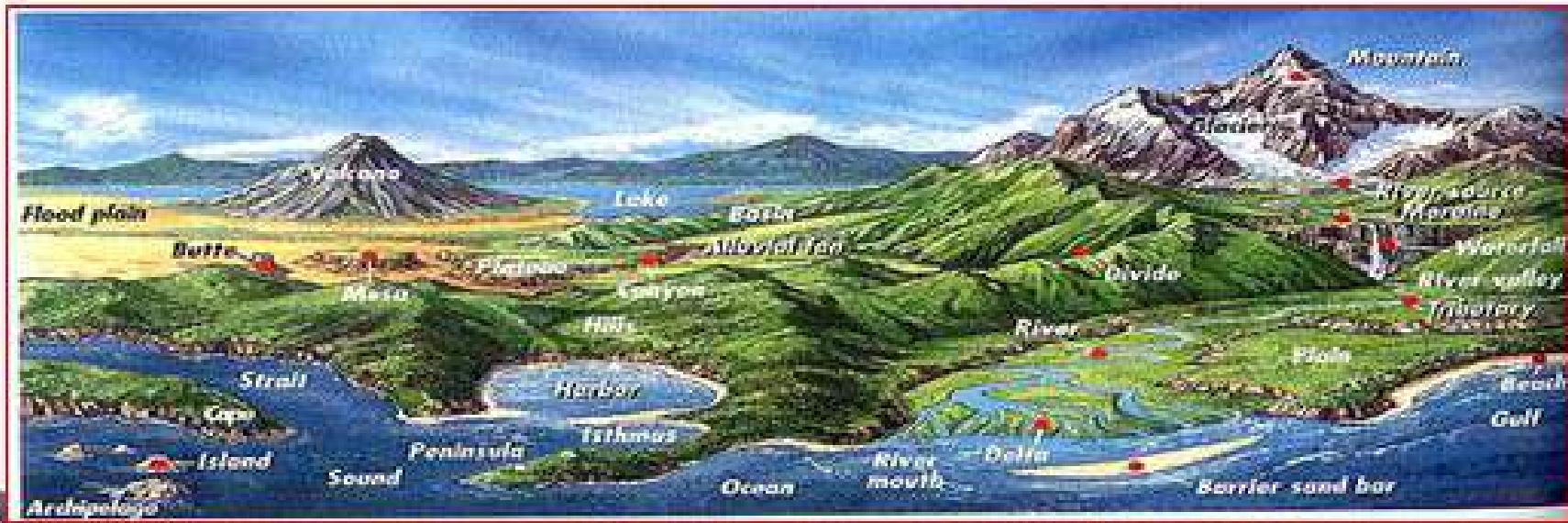
- 这种信息能够用辐射能量的强弱来表征，可记录在感光材料或数字磁带等遥感信息载体上，并能被转换成可视影像。





遥感图像获取

- 一切物体，由于其种类和所处环境条件不同，会具有完全不同的电磁辐射特性。
- 遥感反映了地表资源、环境与灾害等的信息。

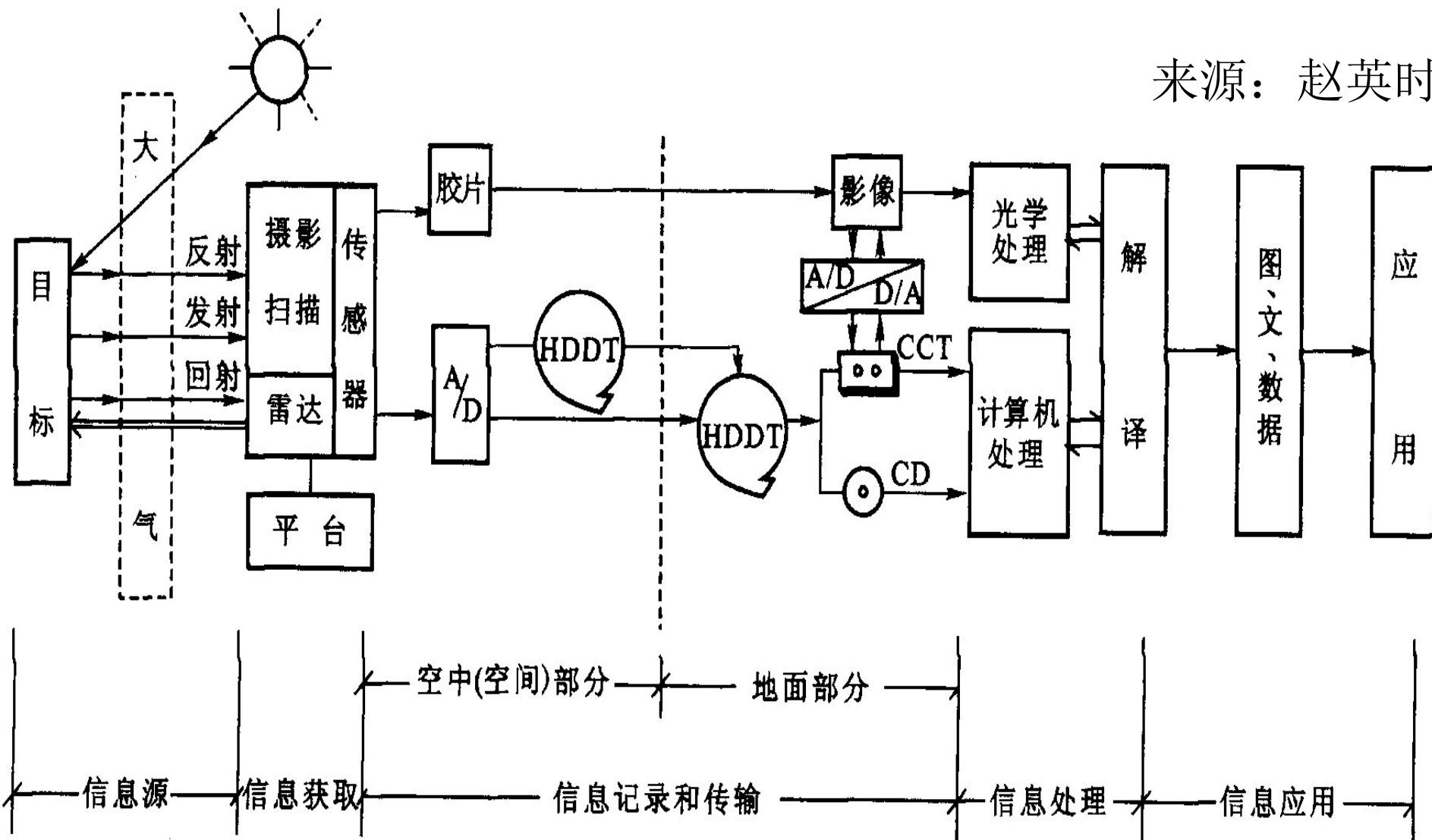


来源互联网：不详



遥感图像获取

3 遥感的过程





4 信息的获取

- **遥感平台：**装载传感器进行遥感探测的运载工具（平台）：
- **传感器：**收集、探测、记录地物电磁波辐射能量的装置。
- **摄影成像：**如分幅式、多光谱
- **扫描成像：**如光机扫描仪、固体自扫描仪、高光谱成像光谱扫描仪
- **微波成像：**如侧视雷达



航天遥感、航空遥感、近地遥感比较

遥感平台
及高度

成像特点

应用特点

航天遥感

位于大气层外的卫星、宇宙飞船等，高度大于**80**千米

扫描：比例尺最小，覆盖率最大，概括性强，具有宏观的特性；多为多波段成像 **雷达**

动态性好，适合对某地区连续观察，周期性好

航空遥感

大气层内飞行的各类飞机、飞艇、气球等，高度小于**20**千米

摄影：比例尺中等，画面清晰，分辨率高，可以对垂直点地物清晰成像；波段数少、**雷达**

动态性差，适合做长周期（几个月及更长）观察

近地遥感

三角架、遥感塔、遥感（船）、建筑物的顶部

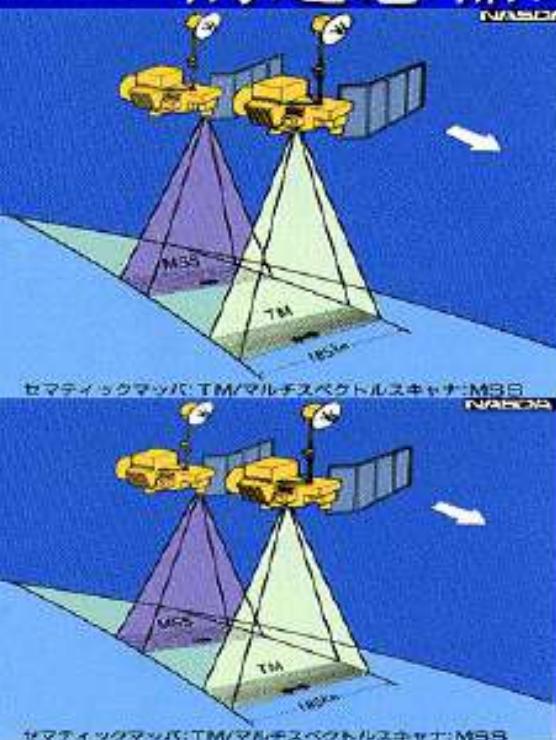
摄影：比例尺最大，覆盖率最小，画面最清晰，多为单一波段成像

灵活机动，费用较低，适合小范围探测



4 信息的获取（中心）

- 地物空间信息主要由搭载在遥感平台上的遥感器来获取



MSS
传感器

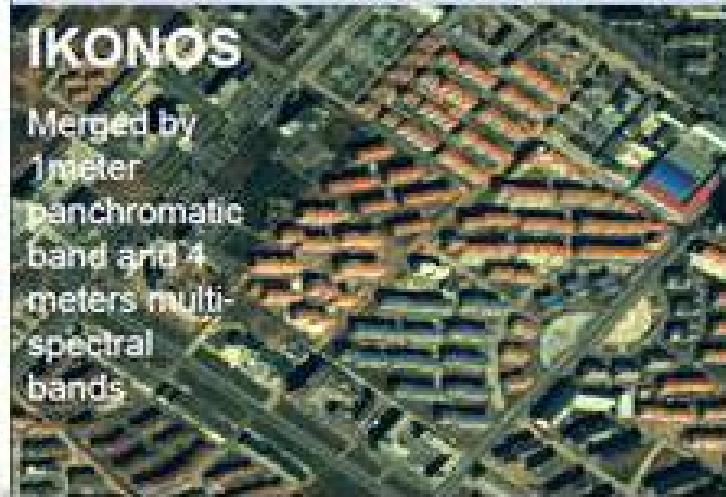
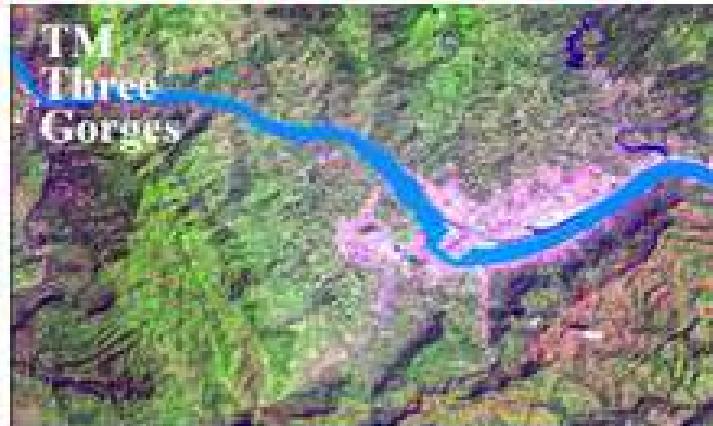
TM
传感器



遥感平台

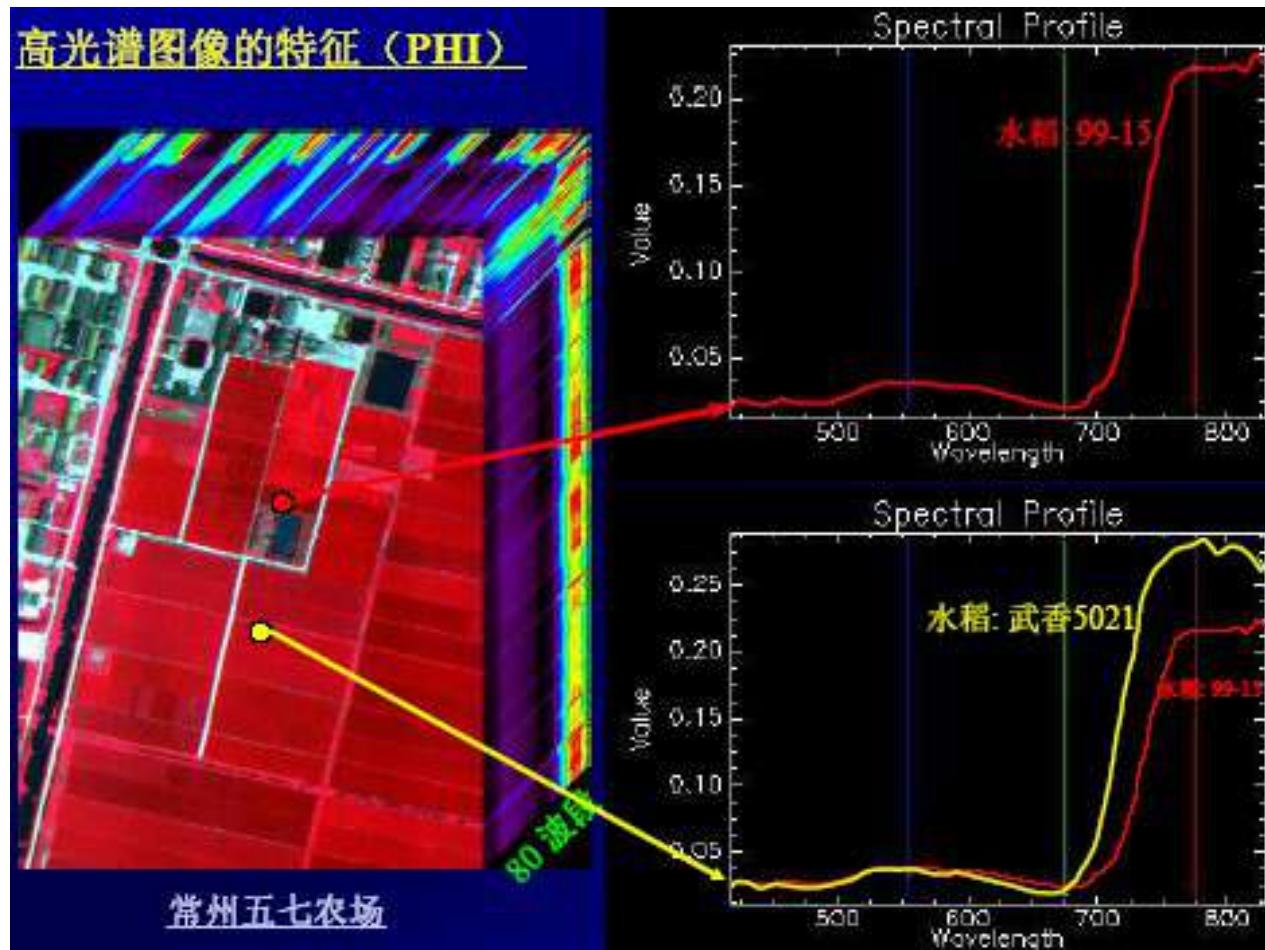


多传感获得遥感数据



来源互联网：不详

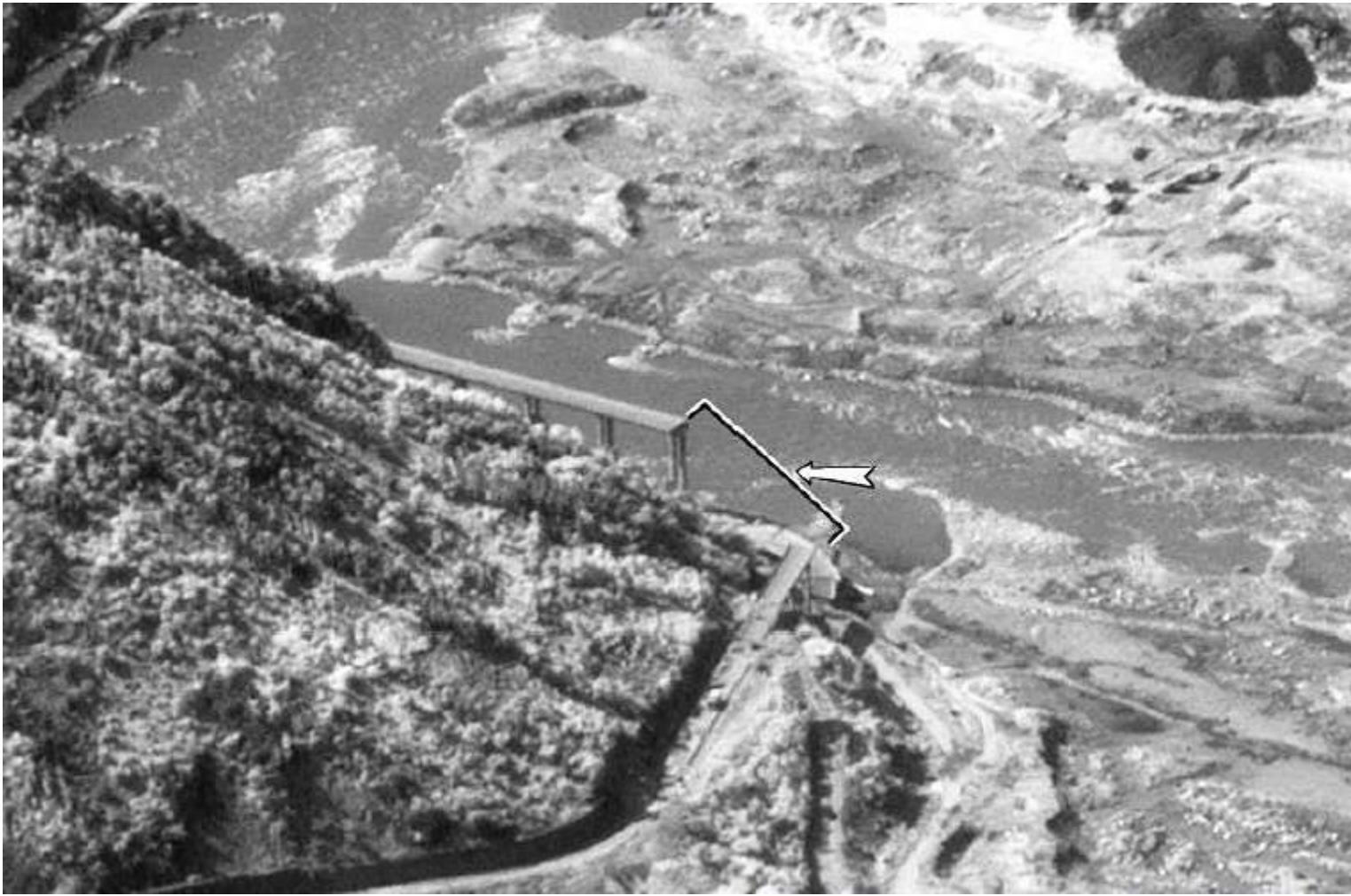




来源互联网：不详



多传感器

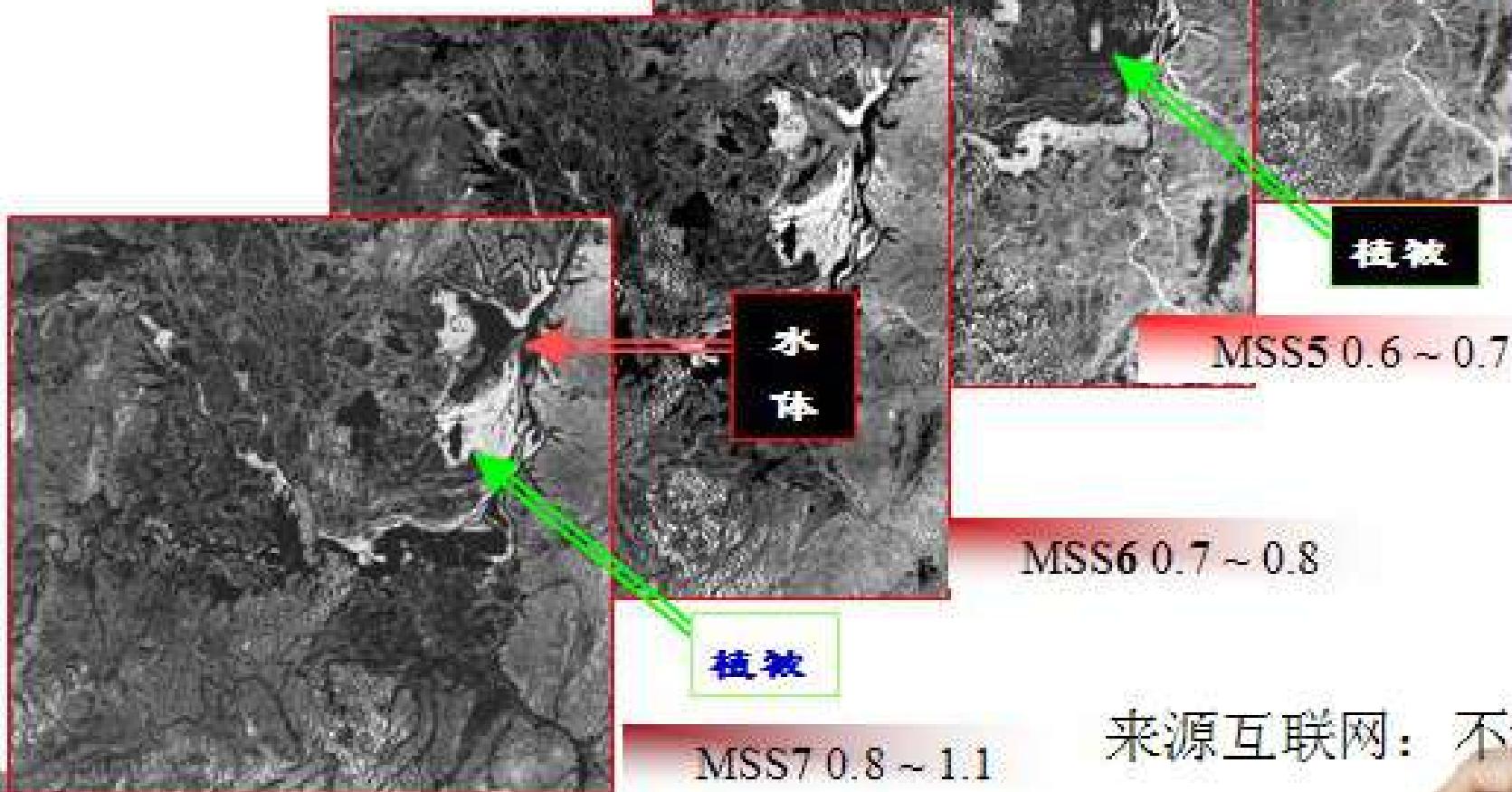


来源互联网：不详



5 数据特点

MSS多波段图像



来源互联网：不详

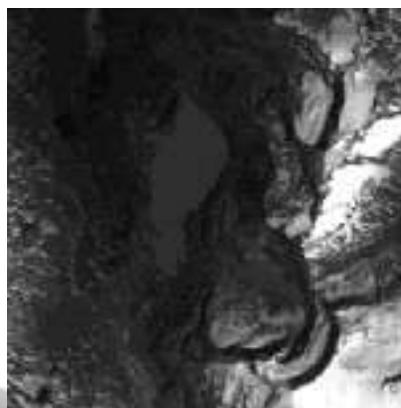


多波段

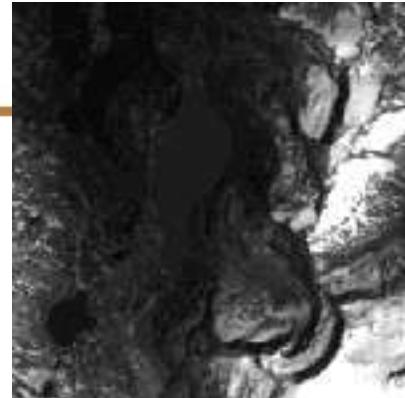
TM 多波段图像



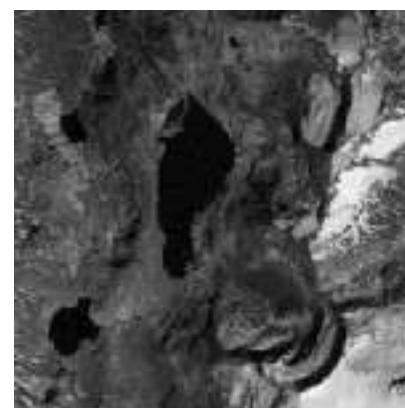
TM1



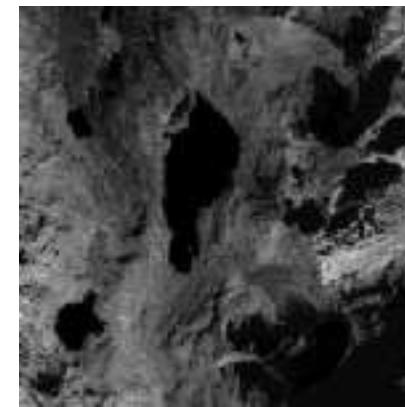
TM2



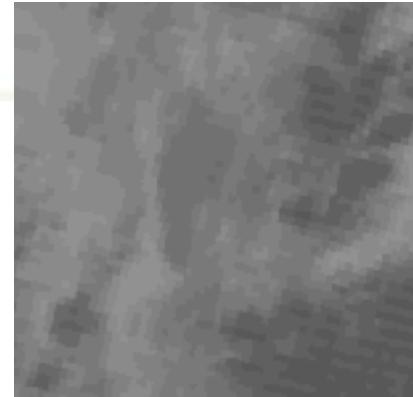
TM3



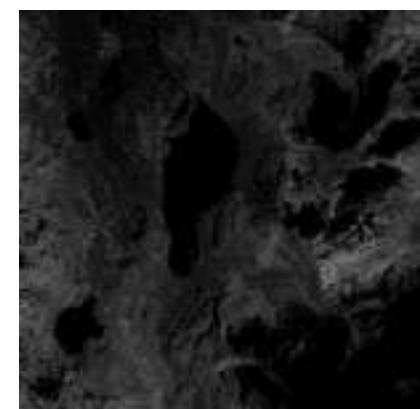
TM4



TM5



TM6



TM7

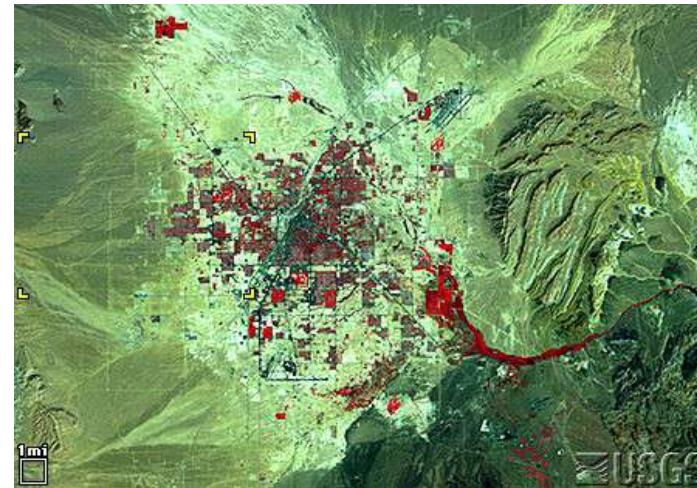
来源互联网：不详



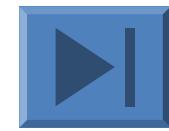
多时相

➤ 多时相性

重复探测，有利于进行动态分析，且成图快。



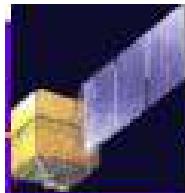
Las Vegas, 1972



来源互联网：不详



多时相



西藏易贡地区特大山体滑坡的遥感监测



2000年5月20日 Landsat-7 卫星影像图。从图中可看出，山体滑坡产生的泥石流堆积成土坝，把河流截断，形成水库型湖泊。

2000. 4. 9, 西藏易贡地区发生特大山体滑坡, 之后又发生了溃坝成灾. 利用多卫星遥感手段有效地监测到山体滑坡与湖泊溃坝情况, 制成卫星影象图, 清晰, 直观地展现了受灾前后的状况.

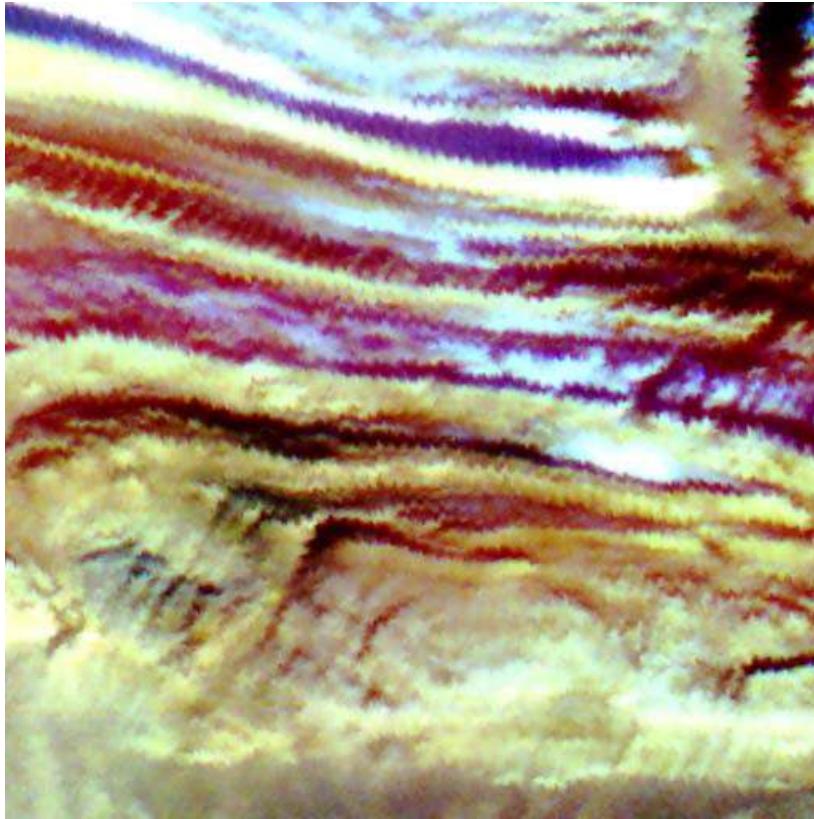
1999年12月20日灾前
的Landsat-5卫星影像图



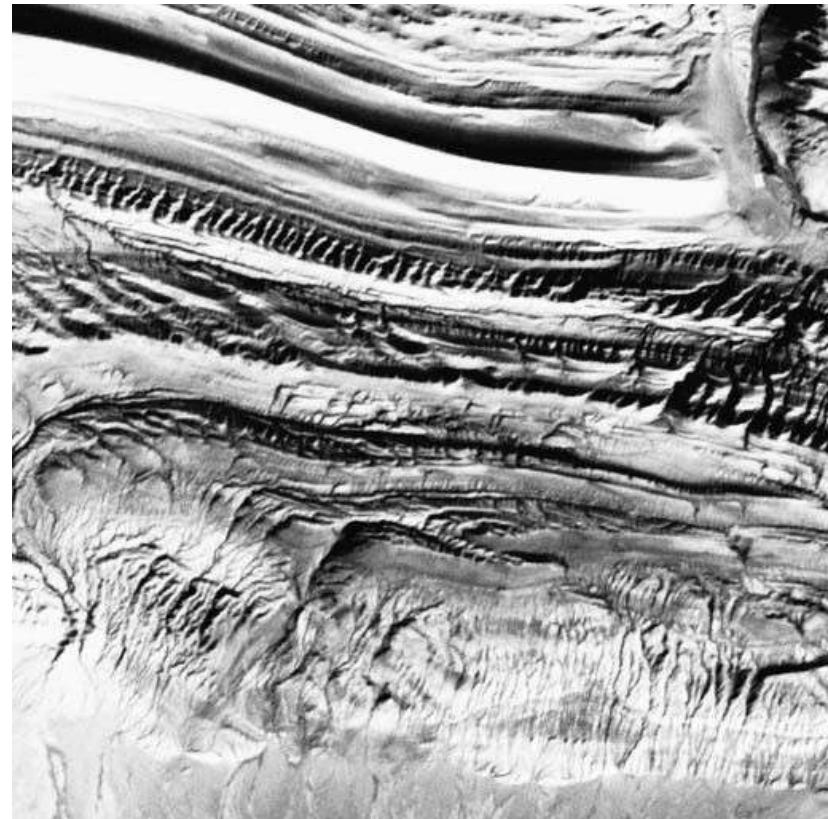
来源互联网：不详



多分辨率



CBERS1 号多光谱



SPOT_{pan}

来源互联网：不详



遥感图像处理

- 遥感图像处理方法包括：
- **一、彩色变换：**
- 单波段彩色变换
- 三波段彩色增强
- HIS 变换
- **二、图像运算：**
- 差值运算
- 比值运算
- **三、多光谱变换**
- **四、图像融合（复合）**
- 不同传感器遥感数据融合
- 不同时相遥感数据融合
- 遥感与非遥感数据融合



一、彩色变换

1、单波段彩色变换（假彩色密度分割）

定义：单波段黑白遥感图像可按亮度分层，对每层赋予不同的色彩，使之成为一幅彩色图像。这种方法又叫密度分割。

处理过程：



图像密度分割原理可以按如下步骤进行：

- (1) 求图像的极大值 d_{\max} 和极小值 d_{\min} ；
- (2) 求图像的密度区间 $\Delta D = d_{\max} - d_{\min} + 1$ ；
- (3) 求分割层的密度差 $\Delta d = \Delta D / n$ ，其中n为需分割的层数；
- (4) 求各层的密度区间；
- (5) 定出各密度层灰度值或颜色。



2、彩色增强（RGB 彩色合成）

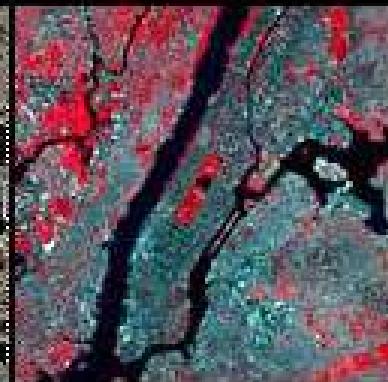
- **定义：**利用计算机将同一地区不同波段的图像存放在不同通道的存储器中，并依照加色法合成原理，分别对各通道的图像赋予红、绿、蓝三种原色，在彩色屏幕上进行叠置，从而构成彩色合成图像。
- **合成方案：**彩色合成图像分为真彩色图像和假彩色图像。方案的选择决定了彩色图像能否突出某些感兴趣信息。
- **方案演示**



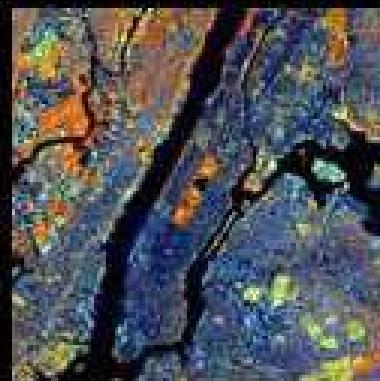
TM bands
3, 2, 1
natural
color image



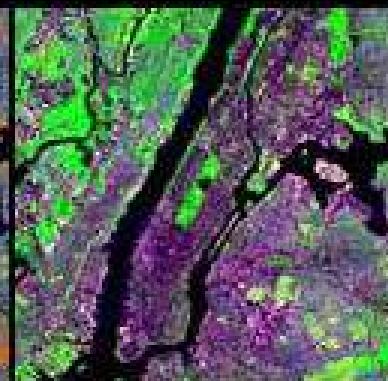
TM bands
4, 3, 2
false-color
infrared



TM bands
4, 5, 3
Increased
definition
of
land-water
boundaries



TM bands
7, 4, 2
Increased
detection
of
moisture in
vegetation
and soils



来源互联网：不详



一、彩色变换

3、HIS 变换

定义：将表色系统 RGB 模式变换为代表色调 H，明度 I 和饱和度 S 构成的 HIS 模式显色模式，这种模式可以用近似颜色立体来定量化；然后再变换为 RGB 系统。

- 作用：**颜色可近似定量表示。

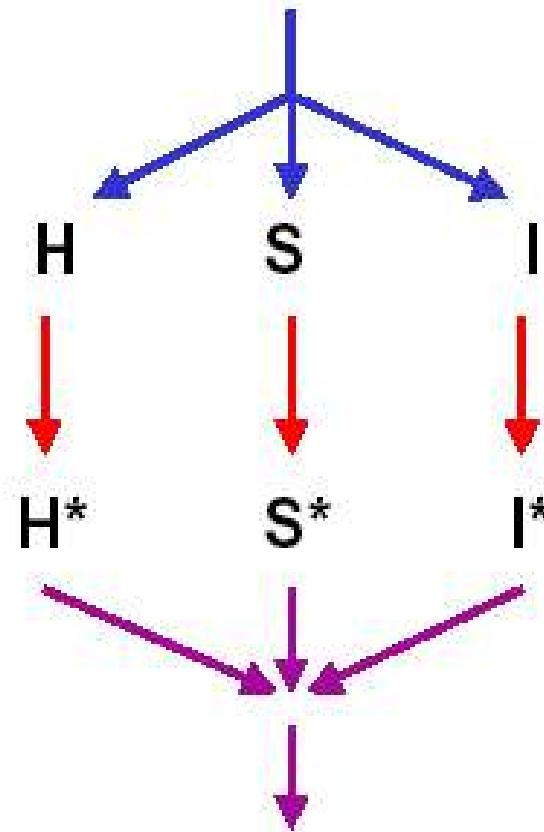


转换

HIS 空间
图像处理

转换

Input image (RGB)



Output image (RGB)

遥感 HIS 变换流程



- 从常用的红、绿、蓝表达方式转换到 HIS 表达方式时，为正变换，有多种变换式。
- 如三角变换
 - $I = \frac{1}{3} (R+G+B)$
 - $H = (G-B)/(3I-3B)$ $S = 1 - B/I$ 当 $B = \min$;
 - $H = (B-R)/(3I-3B)$ $S = 1 - R/I$ 当 $R = \min$;
 - $H = (R-G)/(3I-3B)$ $S = 1 - G/I$ 当 $G = \min$;
其中， $\min = \min[R, G, B]$.
- 将 HIS 模式转换 RGB 模式，为逆变换。





- 将 RGB 模式转换成 HIS 模式，对于定量地表示色彩特性，以及在应用程序中实现两种表达方式的转换具有重要意义。



二、图像运算

- 定义：两幅或多幅单波段影像，完成空间配准后，通过一系列运算，可以实现图像增强，达到提取某些信息或去掉某些不必要信息的目的。
- 原理依据：地物不同波段的光谱差异
 -



1、差值运算：

定义： 两幅同样行、列数的图像，对应像元的亮度值相减就是差值运算。

$$f_d(x, y) = a * [f_1(x, y) - f_2(x, y)] + b \quad a, b \text{ 为经验值}$$

作用：该方法可用于：1) 两个波段影像相减，突出某些地物信息。如红外波段与红波段的差值，可以突出植被信息。

2) 两时相单波段相减，可检测同一地区不同时相的动态变化信息。



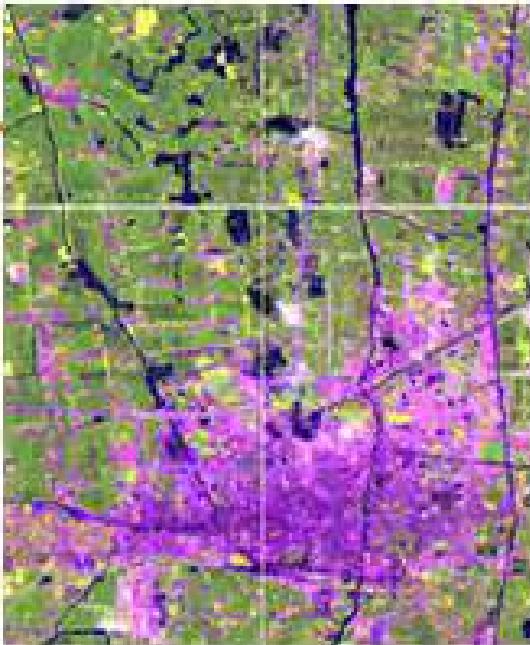
实验

分别用单波段影像（如TM7波段、SPOT_{pan}段），或多波段合成影像比较，监测变化流程如下图：

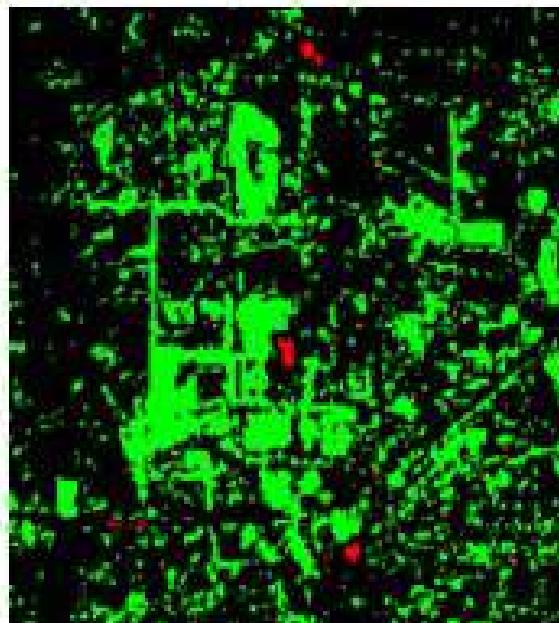




87年合成影像



94年合成影像



87年到94年的变化监测图

阈值可依靠如下经验公式：

$$\text{Average}(X_{ij}) + (\text{or } -) N * \text{STD}$$

Average X_{ij} 为差值图像的均值
STD 为方差，N 取 1 或 2。



2、比值运算：

定义：两幅同样行、列数的图像，对应像元的亮度值相除（除数不为0）就是比值运算。

$$f(x, y) = INT[a \times f1(x, y) / f2(x, y)]$$

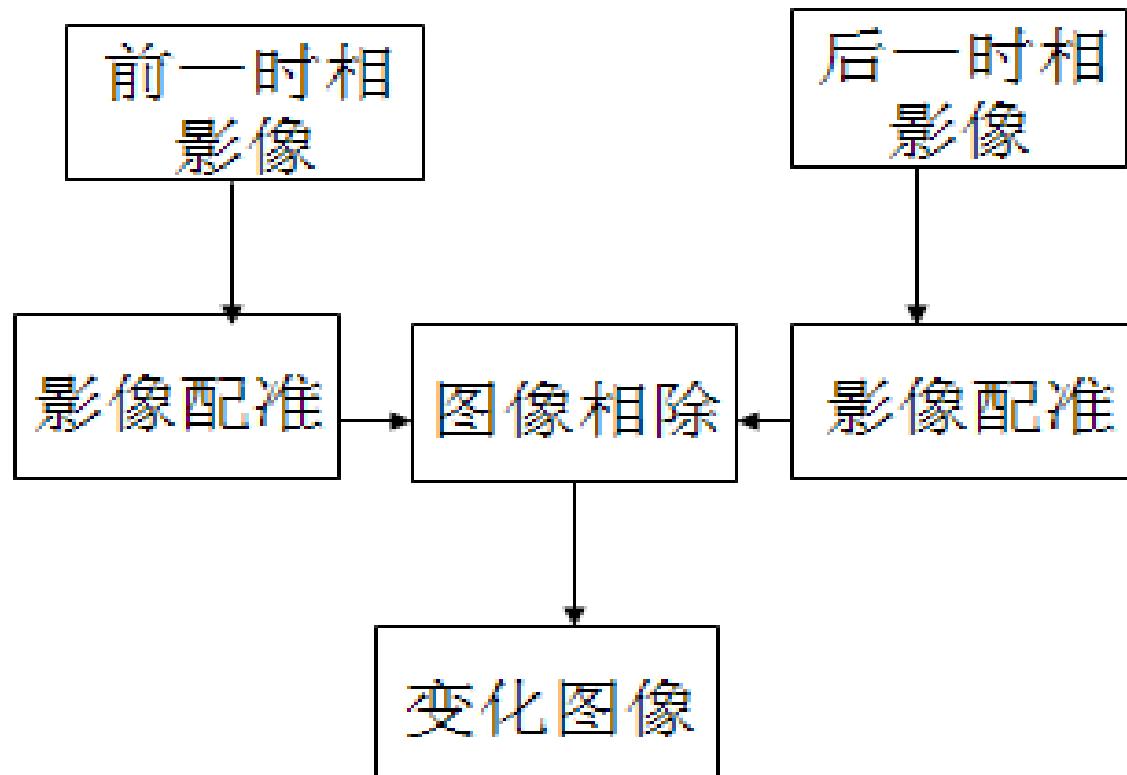
作用：1、突出两个波段的波谱差异，提高对比度，例如植被指数，提取遥感影像中的植被信息；

$$TM4/TM3 \quad NDVI = (TM4 - TM3) / (TM4 + TM3)$$

- 2、消除地形的影响引起阴阳坡亮度的差异；
- 3、不同时相比值可检测变化信息。



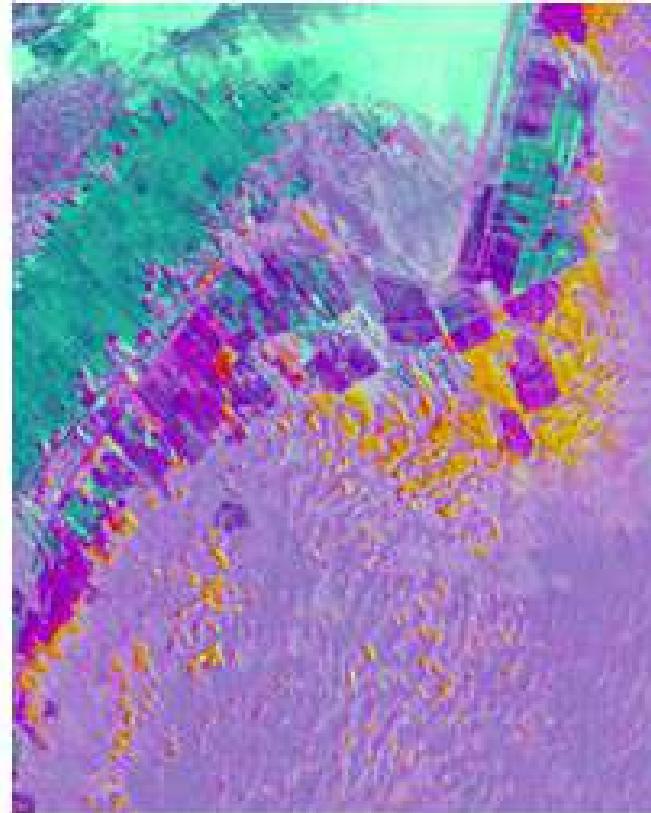
流程：



遥感比值运算流程



实验及分析



$$B3(87a)+B3(94a)+B4(87a)/B4(94a)$$

其中，紫色为新增耕地，橙色为沙漠化土地，变化的特征异常明显。





三、多光谱变换

遥感多光谱影像，特别是陆地卫星的TM等传感器，波段多，信息量大，对图像解译很有价值。

但数据量太大，在图像处理计算时，也常常耗费大量的机时和占据大量的磁盘空间。



实际上，一些波段的遥感数据之间都有不同程度的相关性，存在着数据冗余。

关键：多光谱变换方法可通过函数交换，达到保留主要信息，降低数据量；增强或提取有用信息的目的。

本质：对遥感图像实行线性变换，使多光谱空间的坐标系按一定规律旋转。



1、K-L 变换

• 主成分分析法是一种数学变换的方法，它把给定的一组相关变量通过线性变换转成另一组不相关的变量，这些新的变量按照信息量依次递减的顺序排列。事实上，第一变量具有最大的信息量，称为第一主成分；第二变量的信息量次大，并且和第一变量不相关，称为第二主成分；依次类推， I 个变量就有 I 个主成分。



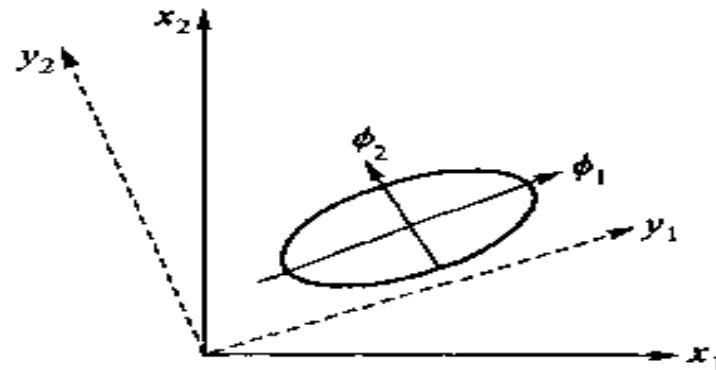
(1). 原理

在它是对某一多光谱图像X. 利用K-L变换矩阵A进行线性组合, 而产生一组新的多光谱图像Y.

$$Y = AX$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1n} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2n} \\ \vdots & & \phi_{ij} & \vdots \\ \phi_{n1} & \phi_{n2} & \cdots & \phi_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix}$$

A的作用是给多波段的象元亮度加权系数, 实现线性变换.



- K-L 变换的特点：
- 变换后的主分量空间与变换前的多光谱空间坐标系相比旋转了一个角度。新坐标系的坐标轴一定指向数据量较大的方向，可实现数据压缩和图像增强。



(2) 特点

- 特点： 1) 数据压缩， 变换后， 第一第二和第三分量包含了绝大部分信息。
- 2) KL变换后的前几个主分量， 信噪比大， 噪声小。突出了主要信息， 达到了增强图像的目的。



(3) 实验

实验1方案为主分量PCI、PC2和PC3的合成，合成影像信息更加丰富，边界更加明显，见实验1图。

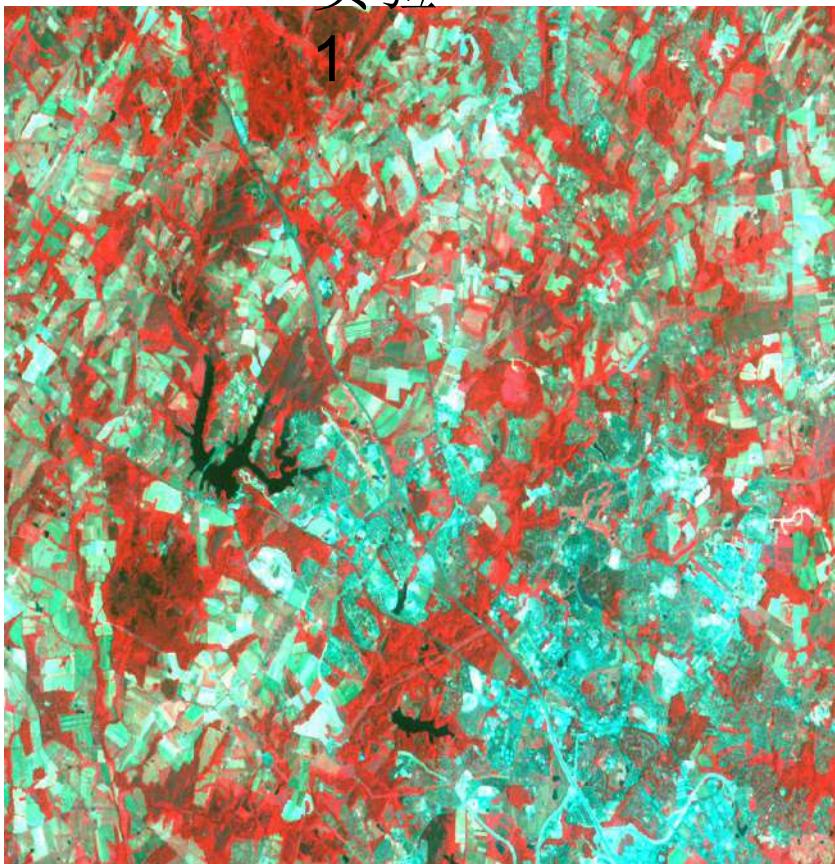
实验2利用两个时相(均为TM5, TM4, TM3)的六个波段经主成分变换后，作PCI、PC2、PC3的合成，主要变化信息以不同色调和纹理突现出来，如果遥感影像检测目标和背景值相差不大，该方法变化信息显示效果较好。

试验表明，该方法发现的图斑与差阈值方法(单波段相减)获得的图斑相对应，见实验2图。

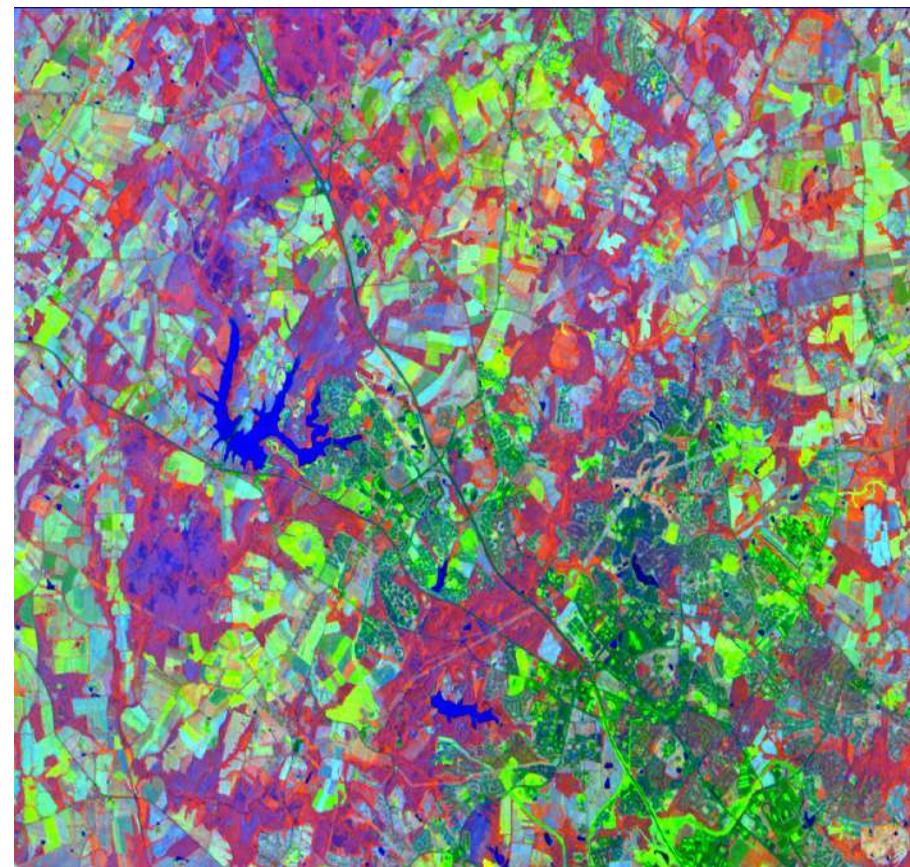


实验

1



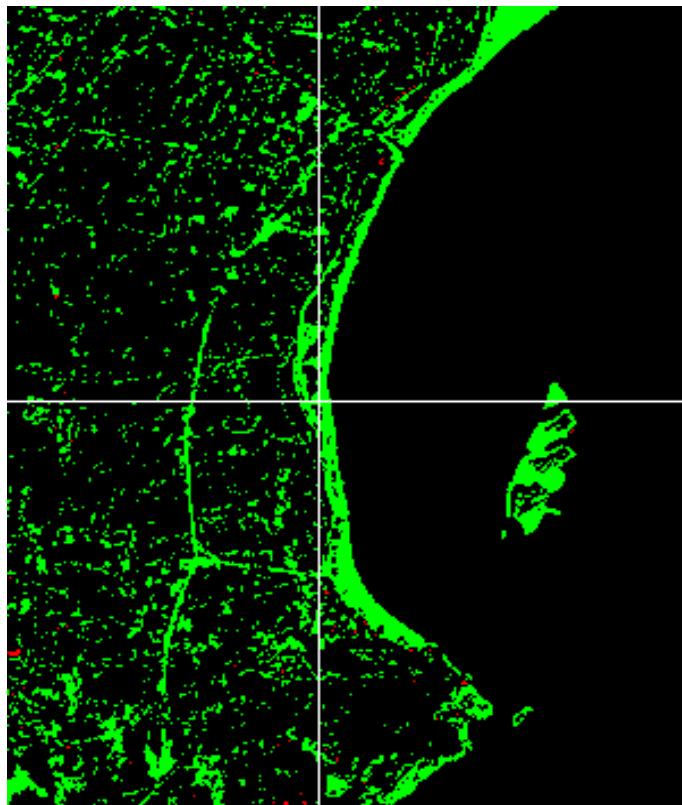
TM4 , 3 , 2 合成图像



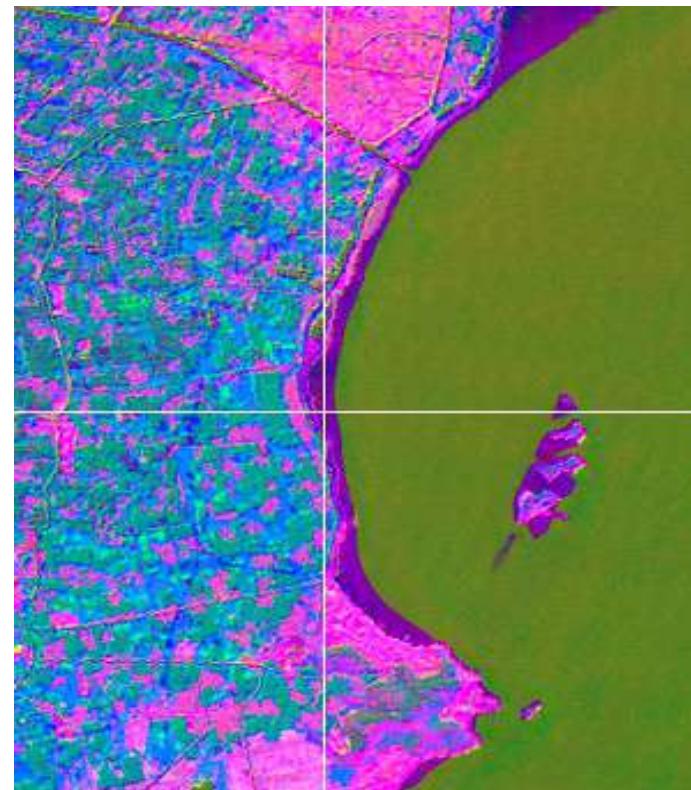
PC1 , 2 , 3 合成图像



实验 2



减法获得的变化图斑



多时相主成分分析影像



2、K-T 变换

- **定义：**

是 kauth-Thomas 变换的简称，也称缨帽变换。是一种坐标空间发生旋转的线性变换，旋转后的坐标轴的方向与地面景物，特别是和植物生长及土壤有密切的关系。

$$Y=BX$$

- **K-T 变换的应用：**
- 主要针对 TM 图像数据和 MSS 数据。对于扩大陆地卫星 TM 影像数据分析在农业环保方面的应用有重要意义。



- 1984年，Crist和Cicone提出TM数据在KT变换中的B值：

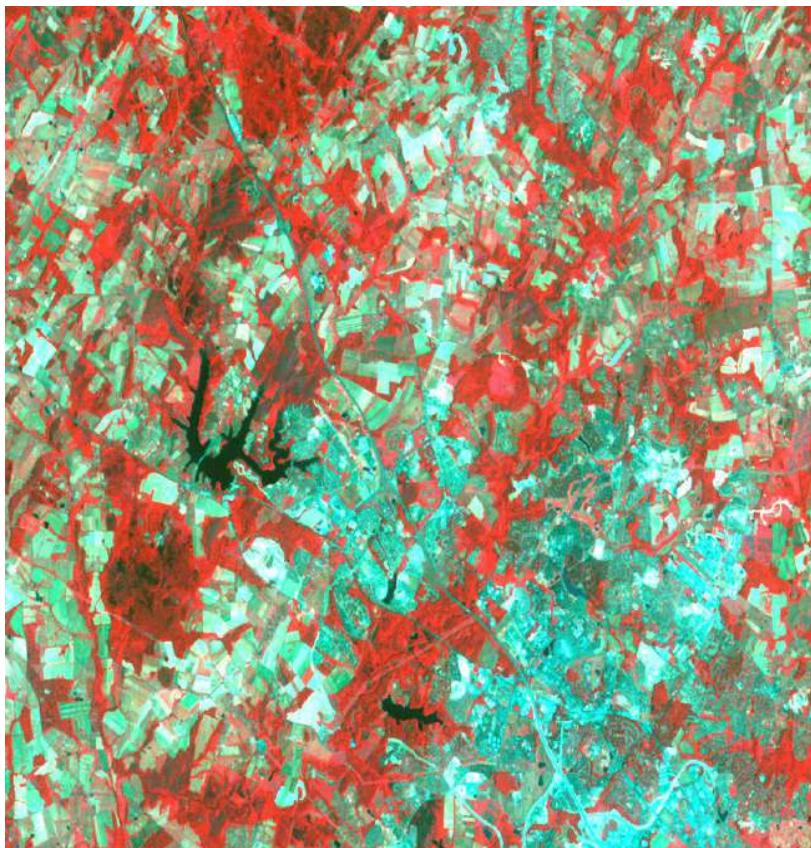
	TM1	2,	3,	4,	5,	7	x1
y1	0.303 7	0.279 3	0.474 3	0.558 5	0.508 2	0.186 3	x1
y2	-0.284 8	-0.243 5	-0.543 6	0.724 3	0.084 0	-0.180 0	x2
y3	0.150 9	0.197 3	0.327 9	0.340 6	-0.711 2	-0.457 2	x3
y4	B = -0.824 2	-0.084 9	0.439 2	-0.058 0	0.201 2	-0.276 8	x4
y5	-0.328 0	-0.054 9	0.107 5	0.185 5	-0.435 7	0.808 5	x5
y6	0.108 4	-0.902 2	0.412 0	0.057 3	-0.025 1	0.023 8	x6



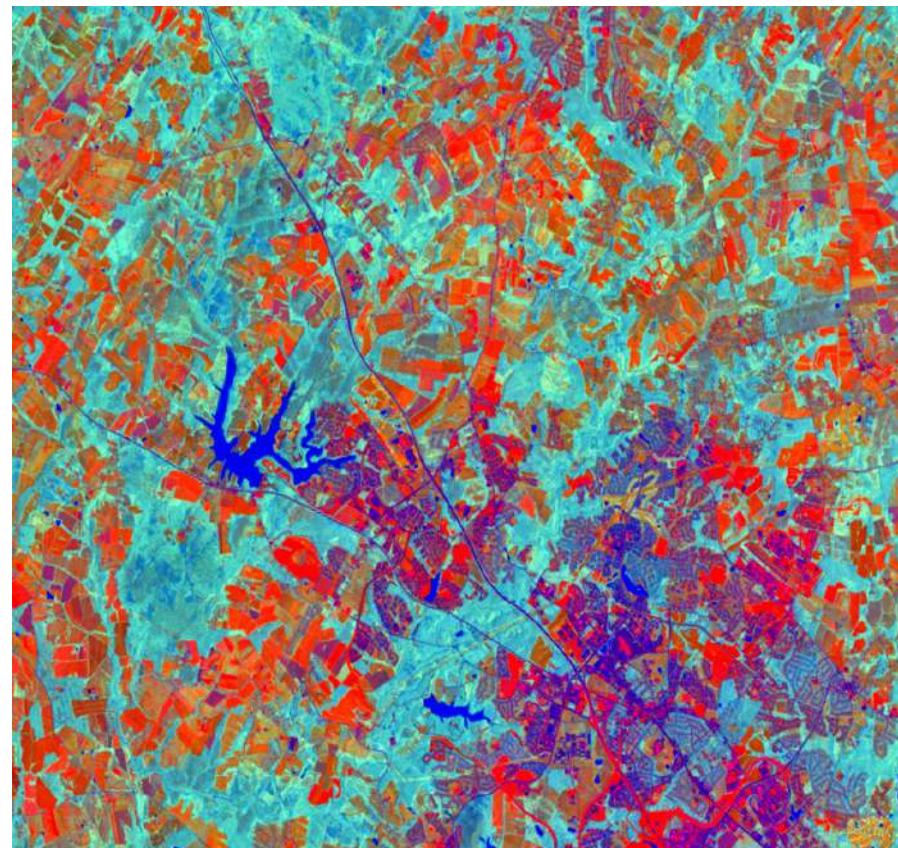
- 这种变换是反映地物特征的几个坐标轴，即亮度，绿度和湿度信息。
- Y1 为亮度，实际是 TM 的 6 个波段加权和，反映出图像的总体反射值；
- Y2 为绿度，反映近红外波段与可见光波段的差值；
- Y3 为湿度，是 1-4 波段与 5， 7 波段的差值，5，7 两波段对土壤湿度和植被湿度最为敏感，易于反映出湿度特征。



实验及分析



TM4,3,2 合成



KT3,2,1 合成



四、多源信息复合

- 定义：
是将多种遥感平台，多时相遥感数据之间以及遥感数据与非遥感数据之间的信息组合匹配的技术。
- 意义：
发挥不同遥感数据源的优势，弥补某种遥感数据的不足，提高遥感数据的应用性；还有利于综合分析和深入理解遥感数据。



四、多源信息复合

1. 不同传感器的遥感数据复合
2. 不同时相的遥感数据复合
3. 遥感与非遥感信息的复合



1、不同传感器的遥感数据复合

(1) 配准：采用几何校正，分别在不同数据源的影像上选取控制点，进行重采样，完成配准。如：TM 数据与 SPOT 数据

(2) 复合：

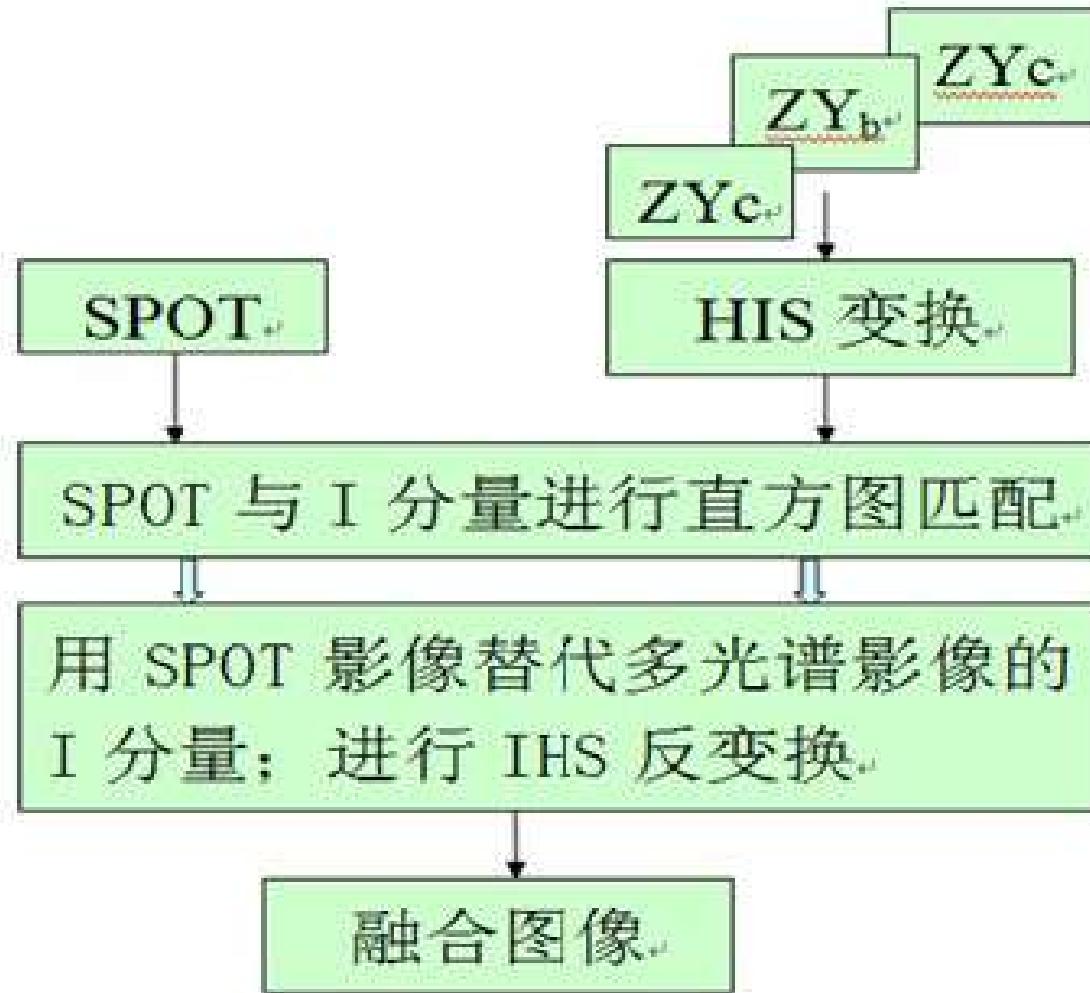
选择融合方法，两者复合既可以提高新图像的空间分辨率，又可以保持较丰富的光谱信息。

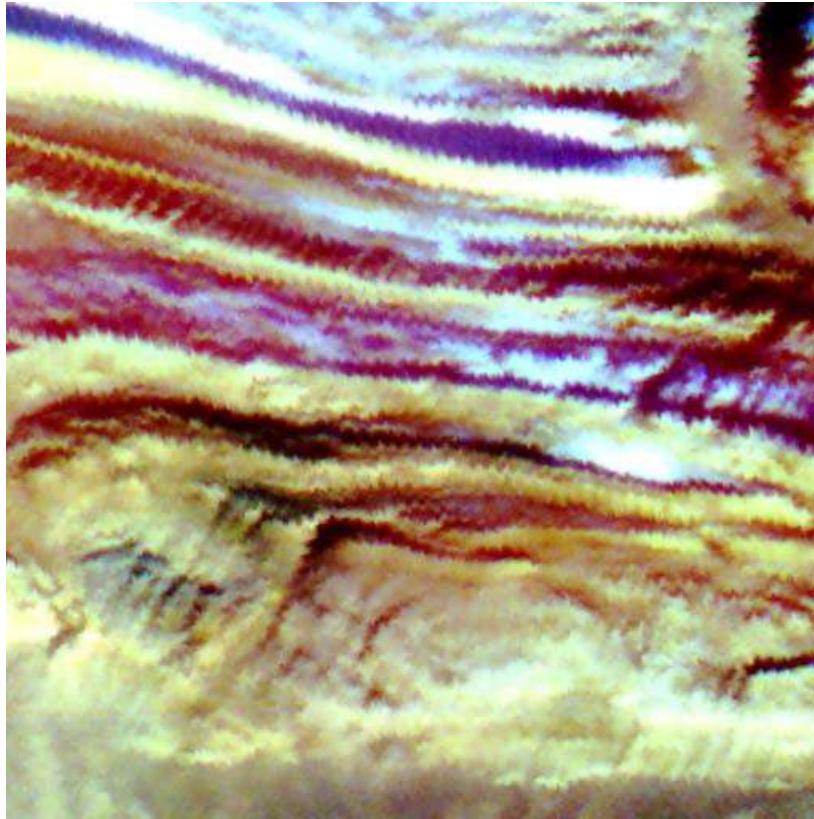


举例：代换法：

- 1)、对TM进行IHS变换；
- 2)、用SPOT全色波段与第I成分进行直方图匹配；
- 3)、用SPOT全色波段替代第I成分；
- 4)、将代换后的所有波段再做一次反变换。

经处理后，如用第四波段（高分辨率数据）的替代I，再经过反变换，再转换为RGB系统，就得到融合图像。





TM

遥感数据源



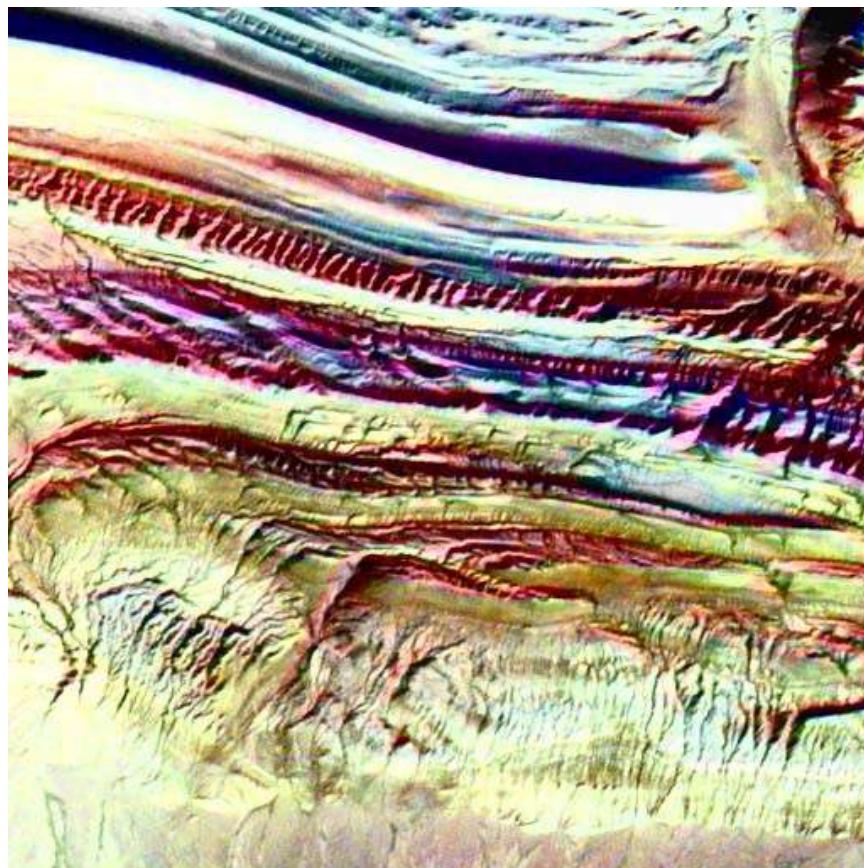
SPOT

来源互联网：不详

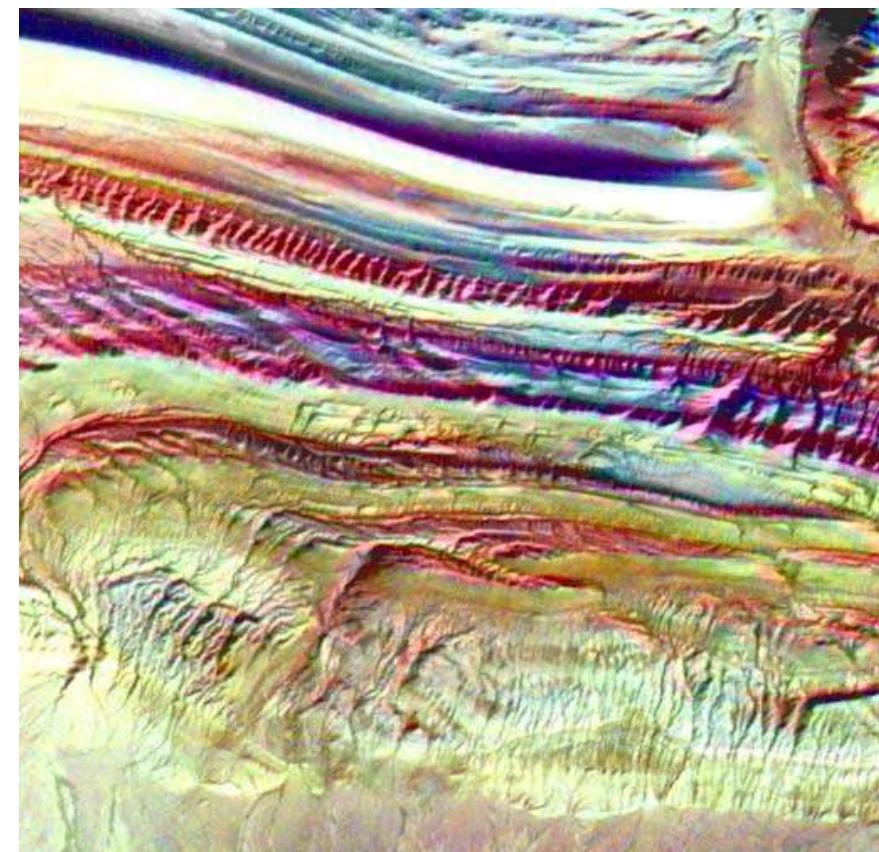




HIS 变换融合结果及分析



不进行直方图匹配



进行直方图匹配

来源互联网：不详





2、不同时相的遥感数据复合

- (1) 配准：利用几何校正的方法进行位置配准。
- (2) 直方图调整：调整成一致的直方图，使图像的亮度趋于一致，便于比较。
- (3) 复合：用于研究时间变化所引起的各种动态变化。

方案：~~SPOT_{pan}~~后时相+TM5, 4, 3前时相

$$S_{R1} = S_{PAN} * \frac{B_R}{B_R + B_G + B_B}$$

$$S_{R2} = S_{PAN} * \frac{B_G}{B_R + B_G + B_B}$$

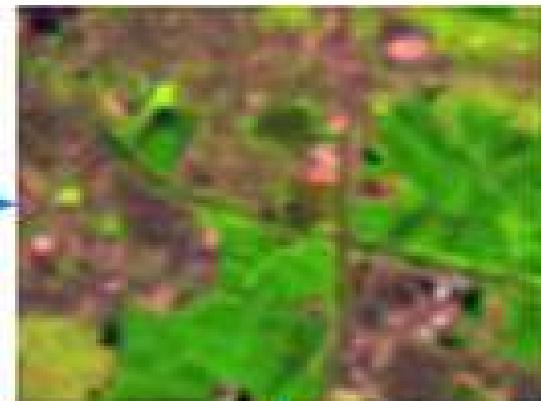
$$S_{R3} = S_{PAN} * \frac{B_B}{B_R + B_G + B_B}$$



前时相 TM



影
像
纠
正



后时相 SPOT



影
像



融 合

亮绿色为
变化的图斑

复合过程

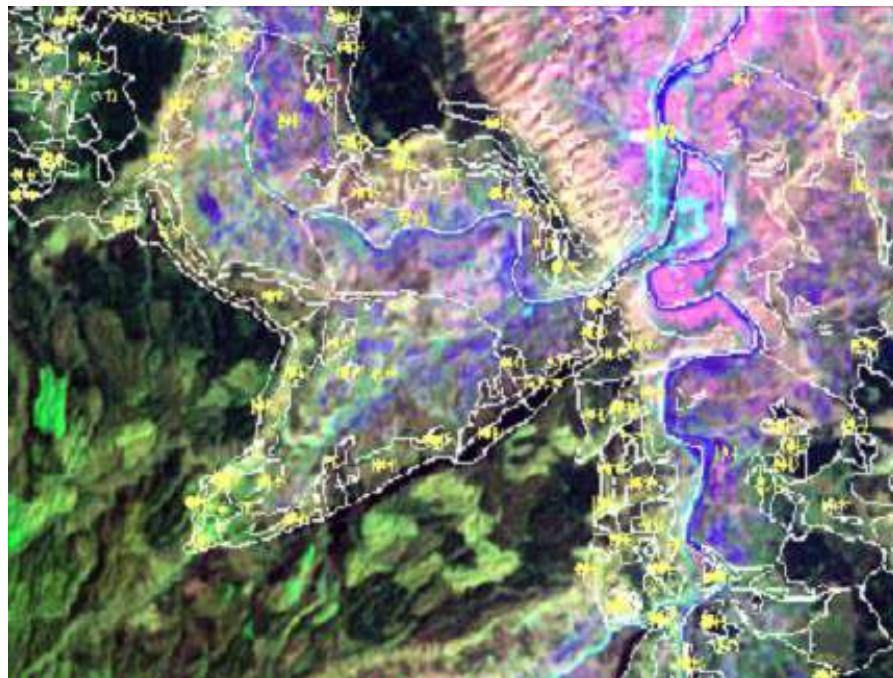


3 、遥感与非遥感信息的复合

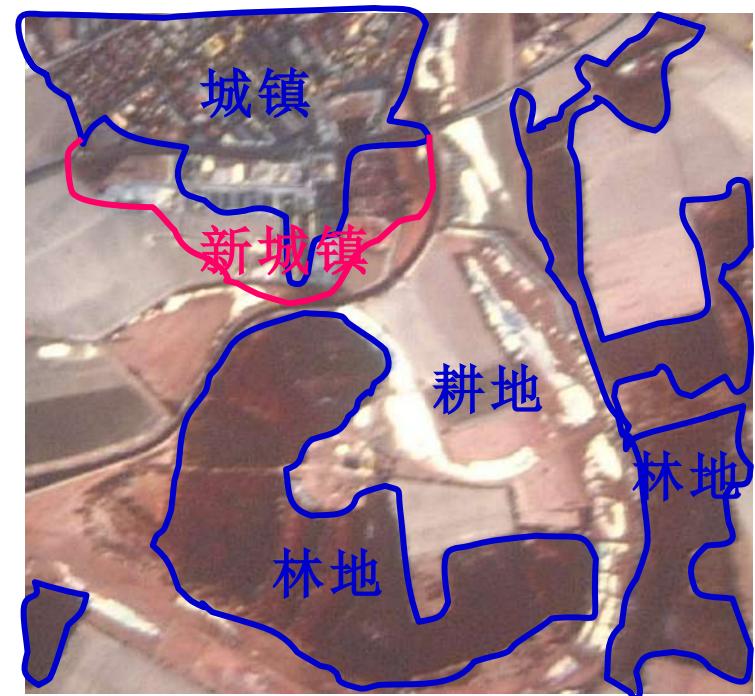
- 一般步骤：
- （1）地理数据的格网化；
- （2）选择最优遥感数据；
- （3）配准复合。



3、遥感与非遥感信息的复合



与地理信息复合



建新图层的复合



小结

本讲介绍遥感的基本概念、遥感的技术系统、遥感数据特点；具体图像处理方法在遥感中的应用，包括彩色增强、图像运算、多光谱变换、图像融合等原理、方法实现以及效果分析等内容。



- 参考书:
- 遥感导论 梅安新
- 遥感应用分析原理与方法 赵英时
- ERDASIMAGINE 遥感图像处理教程 党安荣



谢 谢 !