

# 数字图像处理笔记

陈鸿峥

2020.01\*

## 目录

|                            |           |
|----------------------------|-----------|
| <b>1 概述</b>                | <b>2</b>  |
| <b>2 数字图像基础</b>            | <b>3</b>  |
| 2.1 人类视觉 . . . . .         | 3         |
| 2.2 光和电磁波谱 . . . . .       | 4         |
| 2.3 图像数字化 . . . . .        | 4         |
| 2.4 像素间基本关系 . . . . .      | 6         |
| <b>3 空间域图像增强</b>           | <b>7</b>  |
| 3.1 基本灰度变换 . . . . .       | 7         |
| 3.2 直方图 . . . . .          | 10        |
| 3.3 空间滤波基础 . . . . .       | 13        |
| 3.4 图像锐化 . . . . .         | 15        |
| <b>4 频率域滤波</b>             | <b>18</b> |
| 4.1 傅里叶级数与傅里叶变换 . . . . .  | 18        |
| 4.2 取样函数 . . . . .         | 21        |
| 4.3 二维傅里叶变换 . . . . .      | 22        |
| 4.4 频率域滤波 . . . . .        | 26        |
| 4.5 总结 . . . . .           | 28        |
| <b>5 图像复原</b>              | <b>29</b> |
| 5.1 噪声模型 . . . . .         | 30        |
| 5.2 噪声存在下的空间滤波复原 . . . . . | 31        |
| 5.3 通过频域滤波抑制噪声 . . . . .   | 32        |
| 5.4 线性、位置不变的退化 . . . . .   | 33        |
| 5.5 估计退化函数 . . . . .       | 34        |
| 5.6 逆滤波 . . . . .          | 34        |

\*Build 20200107

|                           |           |
|---------------------------|-----------|
| <b>6 彩色图像处理</b>           | <b>34</b> |
| 6.1 颜色特性 . . . . .        | 34        |
| 6.2 颜色空间 . . . . .        | 35        |
| 6.3 伪彩色处理 . . . . .       | 36        |
| 6.4 全彩色处理 . . . . .       | 36        |
| 6.5 彩色分割 . . . . .        | 37        |
| <b>7 形态学图像处理</b>          | <b>37</b> |
| 7.1 腐蚀与膨胀 . . . . .       | 37        |
| 7.2 开操作与闭操作 . . . . .     | 38        |
| 7.3 击中或击不中变换 . . . . .    | 39        |
| 7.4 基本形态学算法 . . . . .     | 40        |
| 7.5 灰度级形态学 . . . . .      | 41        |
| <b>8 图像分割</b>             | <b>43</b> |
| 8.1 点检测与线检测 . . . . .     | 44        |
| 8.2 边缘检测 . . . . .        | 44        |
| 8.3 边缘连接 . . . . .        | 48        |
| 8.4 边界检测 . . . . .        | 49        |
| 8.5 阈值处理 . . . . .        | 50        |
| 8.6 基于区域的分割 . . . . .     | 52        |
| 8.7 基于形态学分水岭的分割 . . . . . | 52        |
| 8.8 分割中运动的作用 . . . . .    | 54        |
| <b>9 表示与描述</b>            | <b>55</b> |
| 9.1 表示方法 . . . . .        | 55        |
| 9.2 边界描绘子 . . . . .       | 58        |
| 9.3 区域描绘子 . . . . .       | 60        |
| 9.4 主分量描绘 . . . . .       | 63        |

本课程采用书目Rafael C. Gonzalez & Richard E. Woods, *Digital Image Processing (3rd ed)*<sup>1</sup>。

## 1 概述

任何图像本质上就是一个二维函数， $x$ 和 $y$ 是空间坐标，在任何一对空间坐标上的函数值称为该点的强度或灰度。当 $x$ ,  $y$ 和幅值为有限的、离散的数值时，就称这个图像为数字图像。数字图像处理的输入为图像，输出为图像或图像的子集。

基本步骤：图像获取、图像滤波和增强、图像复原、彩色图像处理、小波与分辨率处理、压缩、形态学处理、分割、表示和描述、目标识别。

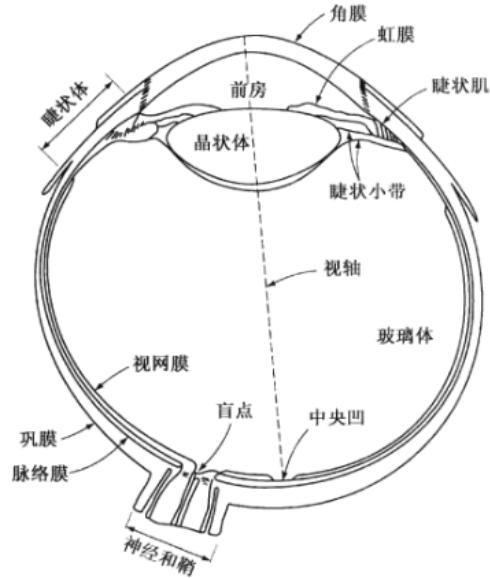
---

<sup>1</sup><http://www.imageprocessingplace.com>

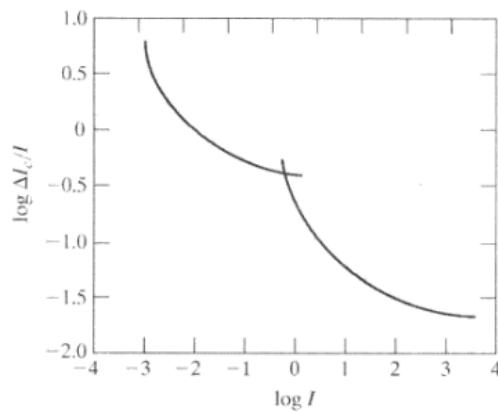
## 2 数字图像基础

### 2.1 人类视觉

- 人的视觉由眼睛中两部分光接收器（感受细胞）组成的：锥状体和杆状体。它们主要分布在视网膜的中间部分，称为中间凹，且对颜色高度敏感。

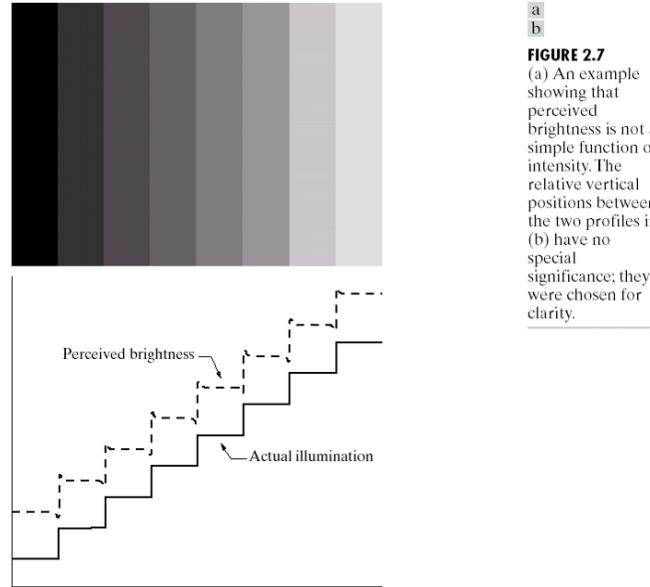


- 成像原理：与光学透镜相似，但适应性强。看远处物体，肌肉会迫使晶状体变得扁平，晶状体的聚焦中心向前移动；物体离眼睛近时，肌肉使晶状体变厚，光心向视网膜成像区域靠近。
- 具体数据：中心凹为视网膜中直径为1.5mm的圆形凹坑，或看为1.5mm×1.5mm的方形传感器阵列。锥状体密度大约为150,000个/mm<sup>2</sup>，数量约337,000个。晶状体中心和视网膜沿视轴的距离/焦距为17mm。
- 韦伯比： $\Delta I_c/I$ ，度量人的眼睛特定的适应级别对亮度变化的辨别力。在低的照明级别，亮度辨别较差（杆状体起作用/微光视觉/暗视觉）。在背景照明增强时，亮度辨别得到明显的改善（锥状体起作用/白昼视觉/亮视觉），即对亮的东西敏感。



- 亮度不是简单的强度函数，下面是两个现象

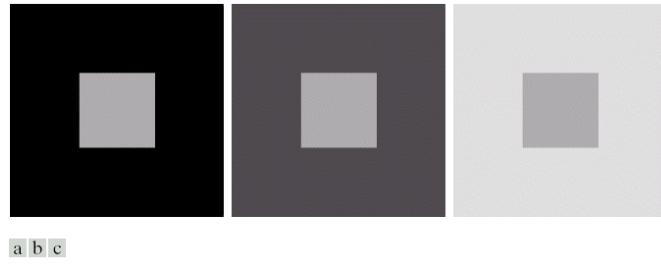
- 视觉系统倾向于不同强度区域边界周围的欠调(undershoot)和过调(overshoot)



**FIGURE 2.7**

(a) An example showing that perceived brightness is not a simple function of intensity. The relative vertical positions between the two profiles in (b) have no special significance; they were chosen for clarity.

- 同时对比现象



**FIGURE 2.8** Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

## 2.2 光和电磁波谱

- 波长 $\lambda$ 与频率 $v$ 关系:  $c = \lambda v$
- 没有颜色的光称为单色光/无色光，其唯一属性是它的强度。因感知单色光的强度从黑色到灰色变化，最后到白色，灰度级通常用来表示单色光的强度。

## 2.3 图像数字化

### 2.3.1 基本概念

用形如 $f(x, y)$ 的二维函数来表示图像， $0 < f(x, y) < \infty$ ，其可以用两个分量来表征

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

其中 $i(x, y) \in \mathbb{R}^+$ 为入射到被观察场景的光源照射总量， $r(x, y) \in (0, 1)$ 为场景中物体所反射的光照总量。

记单色图像在坐标 $(x_0, y_0)$ 处的强度/灰度为 $\ell = f(x_0, y_0) \in [L_{\min}, L_{\max}]$ , 通常令区间为 $[0, L - 1]$ , 其中 $\ell = 0$ 为黑色,  $\ell = L - 1$ 为白色, 其余中间值则为灰度色调。

图像数字化包括两个步骤

- 取样 (时空域): 对坐标值进行数字化, 剖分为像素, 用一个像素值代替这一块,  $f(x, y) \in [0, +\infty)$
- 量化 (光色强度): 对幅值数字化, 变为 $[0, 255]$ 之间的整数

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

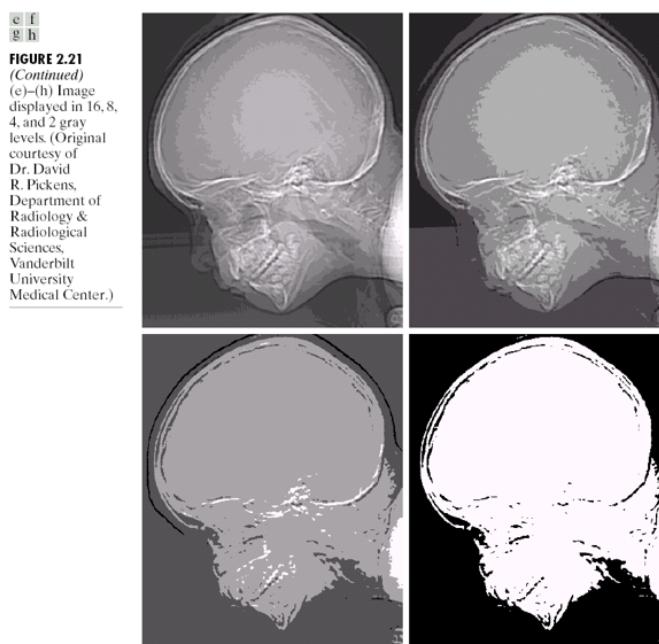
即数字图像包括位置属性 $(x, y)$ 和像素大小 $f(x, y)$ 两个特征, 用上面的二维矩阵表示, 每个元素称为图像的像素, 左上角为零点, 向下为 $x$ 正方向, 向右为 $y$ 正方向。

- 像素越多越精确, 相当于每一块分得越细
- 计算机中字节处理最快, 所以才用8位色(256位)

### 2.3.2 图像分辨率

一副数字图像占用空间 $M \times N \times k$

- 空间分辨率: 图像中可分辨的最小细节, 即图像大小 (行×列)
- 灰度分辨率: 一个像素值单位幅度上包含的灰度级数, 通常是2的整数幂级数 $L = 2^k$ , 如: 用一个byte存一个像素值为256级
  - 人眼对灰度分辨率的敏感程度与图像内容复杂程度有关, 如更偏爱人脸, 而对人群不敏感
  - 当灰度分辨率不够, 会产生伪轮廓



### 2.3.3 图像缩放

将新位置放缩回原图，然后有下列方法

- 最近邻插值：将最邻近像素的灰度将其赋值给新图像中的值
- 双线性(bilinear)插值<sup>2</sup>:  $v(x, y) = ax + by + cxy + d$ , 用周围4个最邻近点进行逼近，4个系数可以用由4个最近邻点的位置方程确定。即设 $Q_{11} = (x_1, y_1), Q_{12} = (x_1, y_2), Q_{21} = (x_2, y_1), Q_{22} = (x_2, y_2)$ , 进而

$$\begin{bmatrix} 1 & x_1 & y_1 & x_1y_1 \\ 1 & x_1 & y_2 & x_1y_2 \\ 1 & x_2 & y_1 & x_2y_1 \\ 1 & x_2 & y_2 & x_2y_2 \end{bmatrix} \begin{bmatrix} d \\ a \\ b \\ c \end{bmatrix} = \begin{bmatrix} f(Q_{11}) \\ f(Q_{12}) \\ f(Q_{21}) \\ f(Q_{22}) \end{bmatrix}$$

- 双三次内插:  $v(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j$

### 2.4 像素间基本关系

**定义 1** (邻域). 坐标为 $(x, y)$ 的像素 $p$ 的

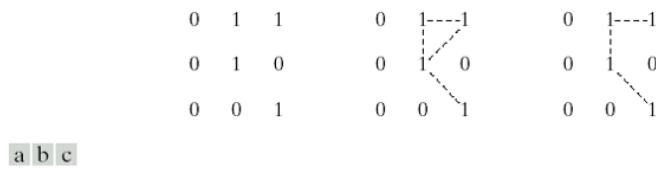
- 4邻域 $N_4(p) = \{(x-1, y), (x+1, y), (x, y-1), (x, y+1)\}$
- 对角相邻像素 $N_D(p) = \{(x-1, y-1), (x+1, y-1), (x-1, y+1), (x+1, y+1)\}$
- 8邻域 $N_8(p) = N_4(p) \cup N_D(p)$

**定义 2** (邻接性). 数字图像的邻接需要满足灰度值的邻接和物理位置的邻接。考虑 $V \subset [0, 255]$ , 有三种邻接性:

- 4邻接:  $q \in N_4(p)$ 且 $q, p \in V$
- 8邻接:  $q \in N_8(p)$ 且 $q, p \in V$
- $m$ 邻接/混合邻接: 要么(i) 4邻接, 要么(ii)  $q \in N_D(p)$ 且 $N_4(p) \cap N_4(q)$ 没有来自 $V$ 中数值的像素

若 $S_1$ 中的某些像素和 $S_2$ 中的某些像素邻接，则这两个集合是邻接的。

注意混合邻接是为了消除8邻接的二义性，见下图。



**FIGURE 2.26** (a) Arrangement of pixels; (b) pixels that are 8-adjacent (shown dashed) to the center pixel; (c)  $m$ -adjacency.

**定义 3** (通路). 若存在 $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$ , 其中 $(x_0, y_0) = (x, y), (x_n, y_n) = (s, t)$ 且 $(x_i, y_i)$ 和 $(x_{i-1}, y_{i-1})$ 是4/8邻接的，则称 $(x, y)$ 到 $(s, t)$ 有4/8通路。

<sup>2</sup>注意这种方法其实不是线性的，计算方法可见<https://www.wikihow.com/Do-a-Double-Linear-Interpolation>

**定义 4 (连通集).** 如果  $S$  中全部像素之间存在一个  $4/8$  通路, 或者说  $S$  仅有一个连通分量, 那么称  $S$  是  $4/8$  连通集。

**定义 5 (距离度量).** 正定性、对称性、三角不等式。图像中主要采用欧氏距离  $D_e$ 、城市街区距离 (*Manhattan*)  $D_4$ 、棋盘距离  $D_8$  (8邻域到中心点的距离都相同)。

### 3 空间域图像增强

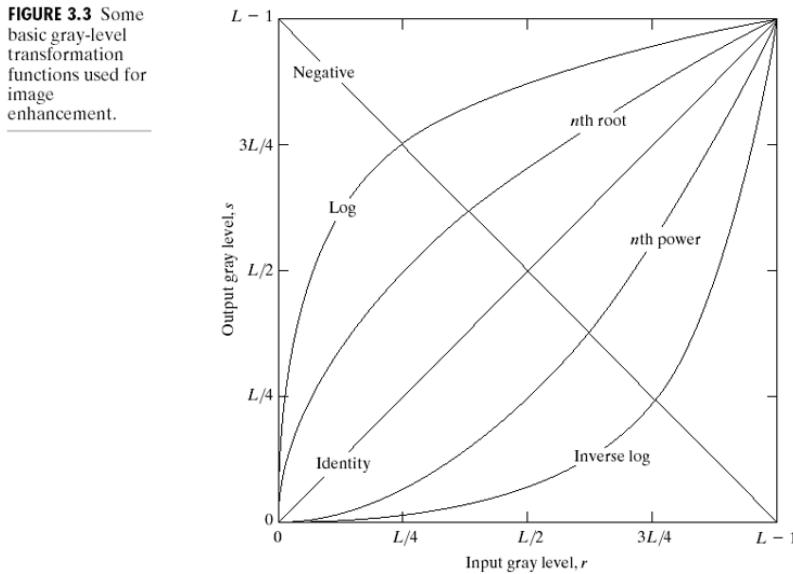
图像增强目的是提高图像在特定应用领域的视觉质量, 包括光滑、锐化、提取边缘、反转、去噪以及各种滤波等。

空间域图像增强的是直接对图像的像素进行操作, 基本关系式可表示如下:

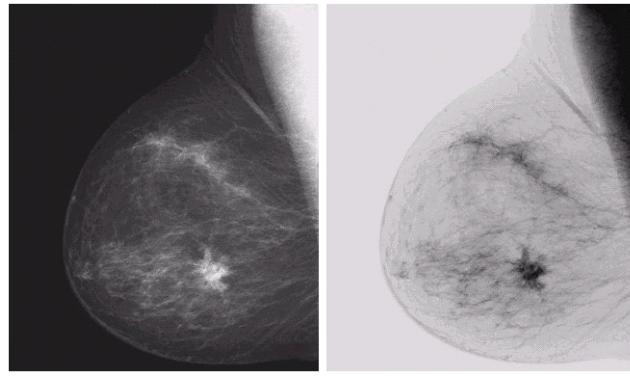
$$g(x, y) = T(f(x, y))$$

#### 3.1 基本灰度变换

- 基本灰度函数: 线性、对数、幂次

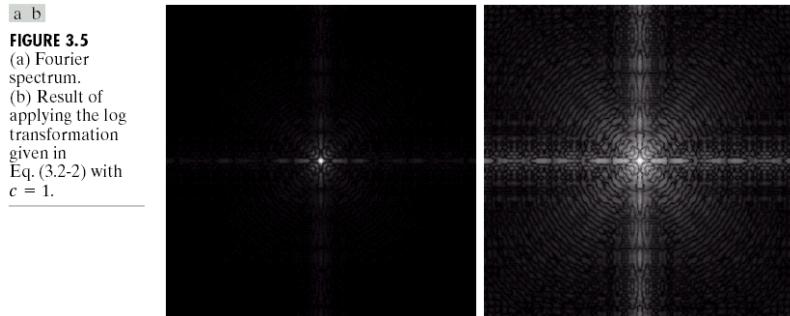


- 图像反转变换:  $s = L - 1 - r$ , 人眼的一个特点就是在背景相对光亮时对灰度层次有较好的分辨能力。



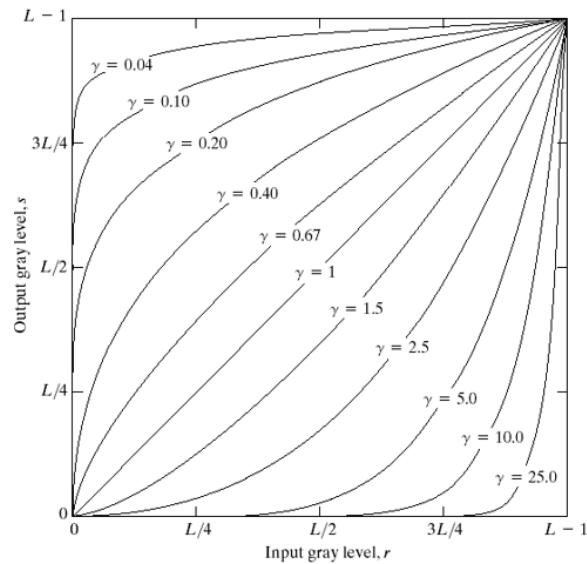
**FIGURE 3.4**  
 (a) Original digital mammogram.  
 (b) Negative image obtained using the negative transformation in Eq. (3.2-1).  
 (Courtesy of G.E. Medical Systems.)

- 对数变换:  $s = c \log(1 + r)$ ,  $c$ 是常数,  $r \geq 0$ , 适合大范围的数据压缩。任何具有对数函数曲线形状的变换都可以完成灰度的压缩和扩展功能。



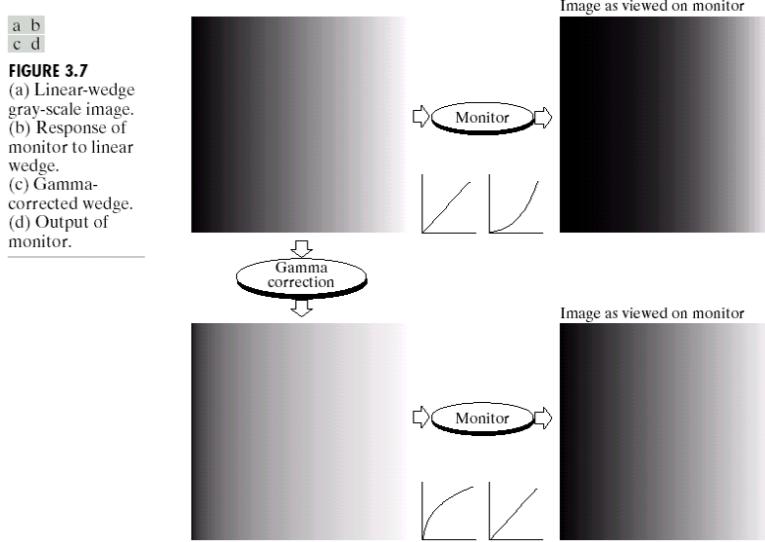
**FIGURE 3.5**  
 (a) Fourier spectrum.  
 (b) Result of applying the log transformation given in Eq. (3.2-2) with  $c = 1$ .

- 幂次变换:  $s = cr^\gamma$ ,  $c$ 和 $\gamma$ 都为正常数



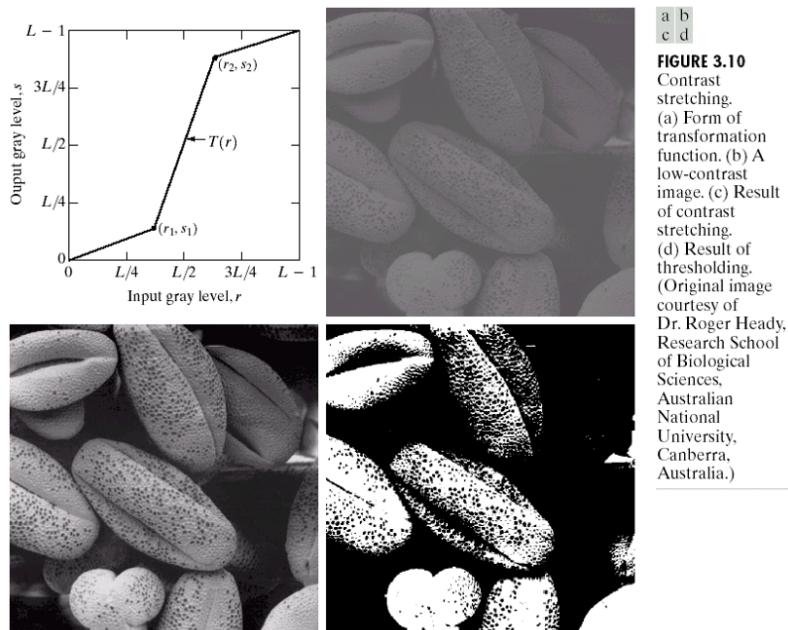
**FIGURE 3.6** Plots of the equation  $s = cr^\gamma$  for various values of  $\gamma$  ( $c = 1$  in all cases).

**伽马校正:** 大量的图像设备如捕捉卡、打印机、数码相机以及显示装置的响应（输出）就对应一个幂函数，通常称这个幂函数的指数为gamma。纠正这个幂次响应的处理称为伽玛校正(gamma correction)。

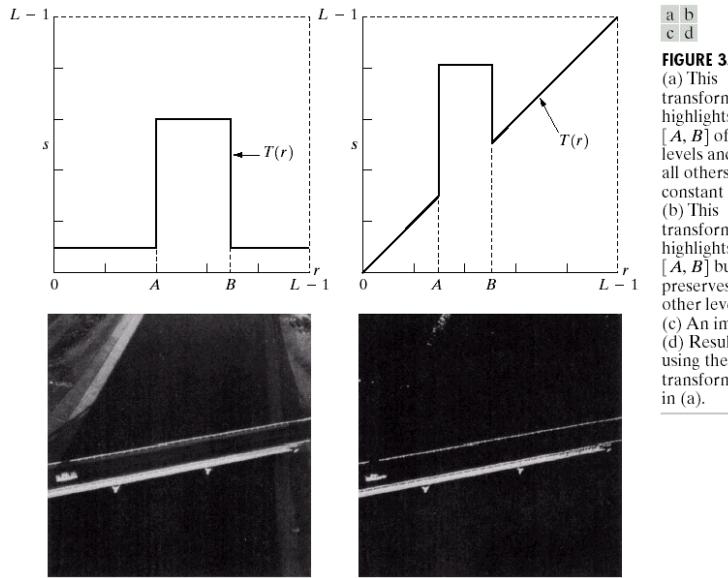


在一般的图像处理软件中，几乎都有伽玛校正的功能。这个功能可用于调整图像的对比度。如果图像偏暗，有些低灰度值的细节被掩盖时，可考虑用指数 $\gamma < 1$ 的伽玛校正（变亮）；反之， $\gamma > 1$ 的校正对那些被“漂白”的细节会起作用（变暗）。

- 分段线性变换

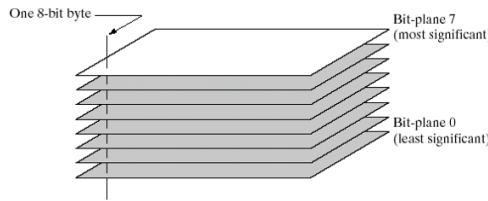


- 灰度切割：在图像中提高特定灰度的亮度



**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).

位图切割：8位灰度图象可以分割成8个位面，每个是一个二值图像（中间切一半）。高位表示了重要的信息，低位给出了不同程度的细节。



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

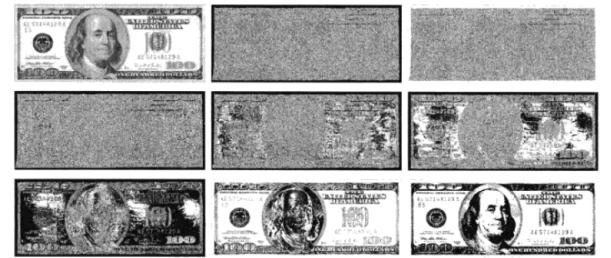


图 3.14 一幅大小为  $500 \times 1192$  像素的 8 比特灰度图像；(b)~(i) 比特平面 1 到 8，其中比特平面 1 对应于最低阶比特。每个比特平面都是一幅二值图像

位图的作用

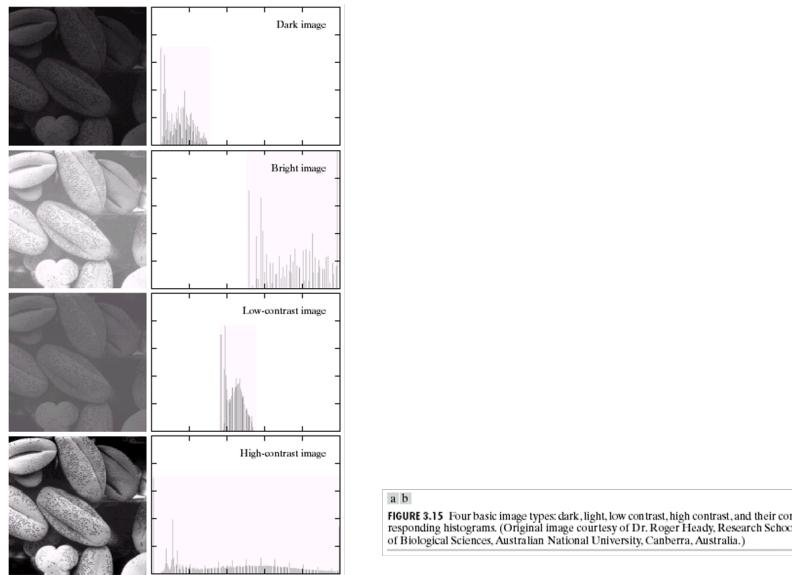
- 信息隐藏：藏在中间位，低位会被丢弃，高位太清楚
- 视频传输：先传高位，再传低位，逐渐清晰

## 3.2 直方图

**定义 6 (直方图)**。灰度级别为  $[0, L - 1]$  (直方图一定从 0 开始!)。数字图像直方图是离散函数  $h(r_k) = n_k$ ，其中  $r_k$  是第  $k$  级灰度， $n_k$  是图像中灰度级为  $r_k$  的像素个数 (频数)。除以总数  $n$  就得到归一化的直方图。

### 3.2.1 直方图均衡化

有亮度差才能看到细节，最好是均衡分布，实现高对比度。



**FIGURE 3.15** Four basic image types: dark, light, low contrast, high contrast, and their corresponding histograms. (Original image courtesy of Dr. Roger Heady, Research School of Biological Sciences, Australian National University, Canberra, Australia.)

直方图均衡化/线性化则是寻求一种变换使得变换后的图像具有尽可能均匀分布的直方图，用于图像增强最大的特点是自动化，有强大的适应性强的功能。通常需要满足下列条件：

- $T(r)$  在  $0 \leq r \leq L - 1$  中为单值且单调递增
- 当  $0 \leq r \leq L - 1$  时， $0 \leq T(r) \leq L - 1$

步骤如下：

1. 概率  $p_r(r_k) = n_k/n, k = 0, 1, 2, \dots, L - 1$

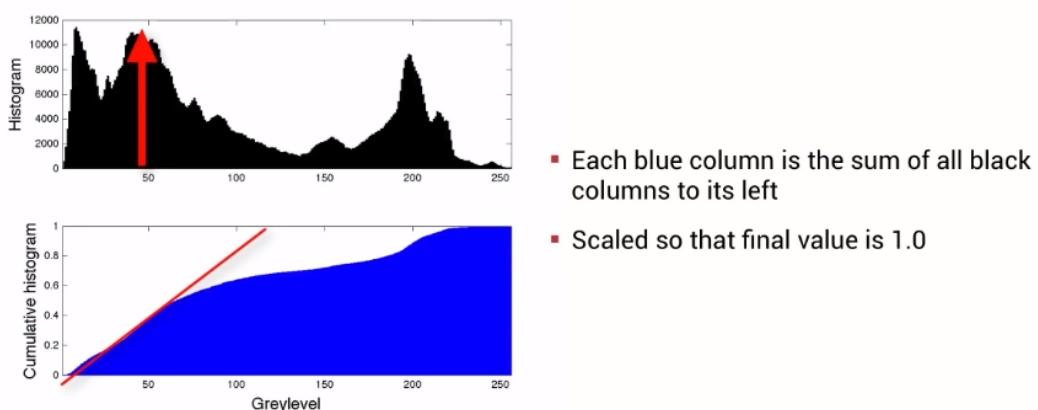
2. 累计分布函数(PDF)

$$P(r_k) = \sum_{j=0}^k p_r(r_j) = \sum_{j=0}^k \frac{n_j}{n}, k = 0, 1, 2, \dots, L - 1$$

3. 变换函数

$$s_k = T(r_k) = (L - 1) \cdot \sum_{j=0}^k \frac{n_j}{n}, k = 0, 1, 2, \dots, L - 1$$

4. 将  $s_k$  四舍五入转换为标准灰度级别，如有相同  $[s_k]$  则合并



分析. 若 $r$ 为离散型随机变量,  $T(r)$ 为单调递增函数 ( $T^{-1}$ 存在且单调), 且 $s = T(r)$ , 则由概率论有

$$p_s(s) = p_r(r) \left. \frac{\partial r}{\partial s} \right|_{r=T^{-1}(s)} = p_s(T^{-1}(s)) |T^{-1}(s)|$$

考虑变换函数为 $r$ 的累积分布函数(CDF)

$$s = T(r) = \int_0^r p_r(\omega) d\omega$$

对 $s$ 求导并代入上面的式子可得 $p_s(s) = 1$ , 故实现均衡。算法实际上就是求累积分布函数, 使得原本亮度小的像素能够映射到亮度大的空间。

### 3.2.2 直方图匹配

当两幅图像比对时, 通常要使其直方图形式一致(如不同光照条件下的同一场景)。先是做空间归一化(伸缩、旋转), 然后再做像素的归一化。

做法是使两幅图像均衡化后的结果相同, 即

$$\begin{aligned} s &= T(r) = \int_0^r p_r(w) dw \\ s &= G(z) = \int_0^z p_z(t) dt \\ \implies z &= G^{-1}(s) = G^{-1}(T(r)) \end{aligned}$$

具体步骤如下:

1. 计算出原图的直方图

$$s_k = T(r_k) = (L-1) \sum_{j=0}^k \frac{n_j}{n}$$

2. 给出输出图像期望的直方图 $p_z(z)$ , 并令

$$v_k = G(z_k) = (L-1) \sum_{i=0}^k p_z(z_i) = s_k$$

3. 寻找区间 $[0, L-1]$ 的最小整数 $\hat{z}$ , 使得

$$G(\hat{z}) - s_k \geq 0$$

即由 $k$ 映射到 $\hat{z}$

**定义 7 (n阶矩).** 设 $p_r(r_k) = n_k/n$ , 则 $r$ 的第 $n$ 阶中心矩为

$$\mu_r(r_k) = \sum_{i=0}^{L-1} (r_i - \bar{r})^n p(r_k)$$

其中 $\bar{r} = \sum_{i=1}^{L-1} r_i p(r_i)$ 为 $r$ 的平均值。特别地, 当 $n=2$ 时为方差。

### 3.3 空间滤波基础

滤波的概念来自信号处理中的傅里叶变换，空间滤波指的是直接对图像像素进行处理的操作。滤波器(filter)有时也叫掩模(mask)、核(kernel)、模板(template)或窗口(window)。

#### 3.3.1 线性滤波

空间域线性滤波基本公式：

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t)$$

常见的情况是， $a = b$  为奇数，如 1、3、5。

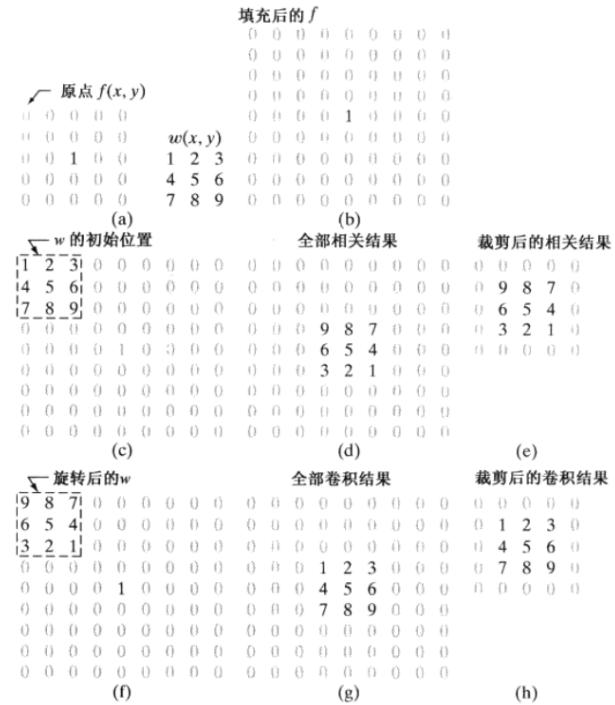
**定义 8** (相关与卷积). 一个  $m \times n$  的滤波器  $w(x, y)$  与  $M \times N$  图像  $f(x, y)$  的相关操作定义为

$$w(x, y) \circ f(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t)$$

一个  $m \times n$  的滤波器  $w(x, y)$  与  $M \times N$  图像  $f(x, y)$  的卷积定义为

$$w(x, y) * f(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x - s, y - t)$$

其中右侧等号表示将  $f$  旋转  $180^\circ$  (或先沿  $x$  轴翻折，再沿  $y$  轴翻折)



空间滤波对边界的处理方法

- 重复边缘值
- 卷绕输入图像
- 补零（最常用）
- 勿略

常用的平滑滤波器（左侧是均值滤波）

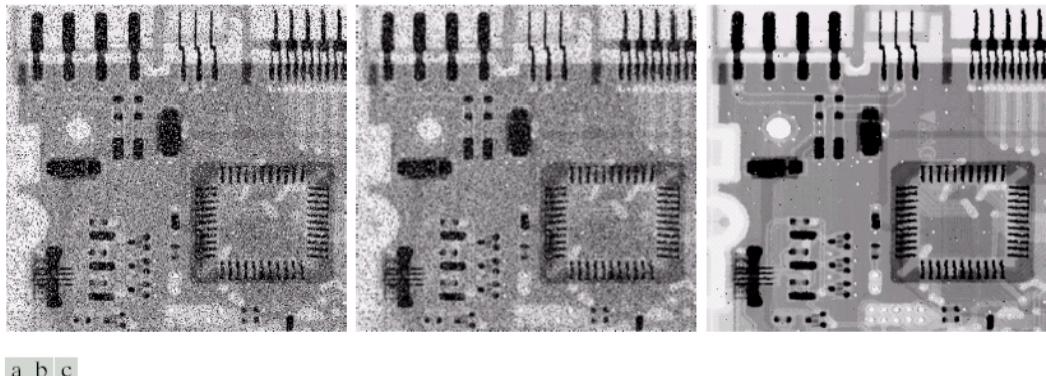
$$g(x, y) = \frac{\sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t)}{\sum_{s=-a}^a \sum_{t=-b}^b w(s, t)}$$

|     |     |     |      |     |      |
|-----|-----|-----|------|-----|------|
| 1/9 | 1/9 | 1/9 | 1/16 | 1/8 | 1/16 |
| 1/9 | 1/9 | 1/9 | 1/8  | 1/4 | 1/8  |
| 1/9 | 1/9 | 1/9 | 1/16 | 1/8 | 1/16 |

### 3.3.2 非线性滤波

排序统计滤波器是一种非线性的、非卷积滤波器。排序统计滤波器在滤波器包围的像素范围内排序，然后由统计排序结果决定的值代替中心像素的值。按排序输出的位置分，可分为：中值滤波、最大值滤波、最小值滤波。

中值滤波比均值滤波更适合做椒盐噪声<sup>3</sup>的去除，因为噪声总是最大最小，故做中值滤波容易去除。



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

同理除了中值滤波，也可以构造第 $X$ 百分点的滤波器。通常 $X < 50$ 图像趋于变暗， $X > 50$ 图像趋于变亮。

<sup>3</sup>特别大或特别小的噪声

## 3.4 图像锐化

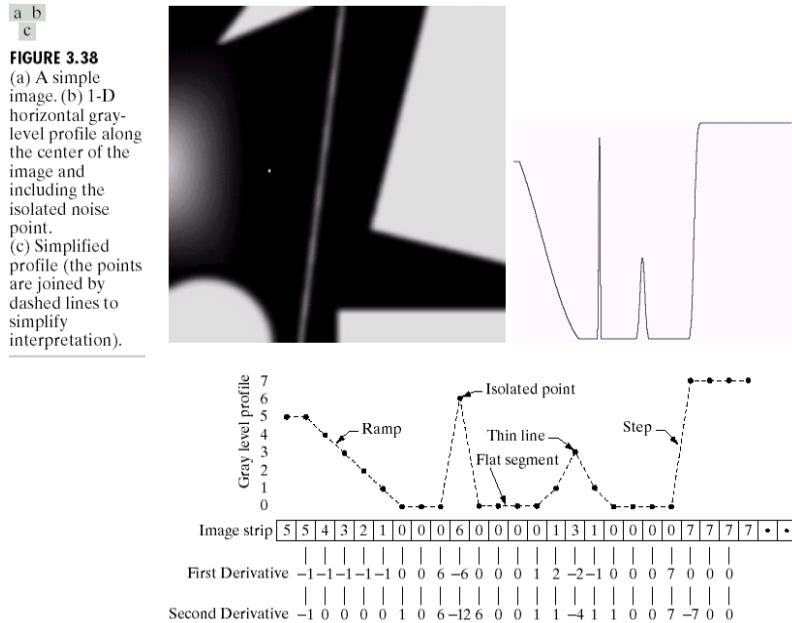
### 3.4.1 图像微分

积分运算可以做平滑，微分运算可以做锐化！ 锐化的目的即突出图像中的细节或者增强被模糊的细节。

考虑离散情况

$$\begin{cases} \frac{\partial f}{\partial x} &= f(x+1) - f(x) \\ \frac{\partial^2 f}{\partial x^2} &= f(x+1) + f(x-1) - 2f(x) \end{cases}$$

斜边缘、 $\delta$ /冲击边缘、阶梯边缘



- 一阶微分产生较“宽”的边界，二阶微分产生较“细”的边界
- 二阶微分处理对细节有较强的响应，如细线和孤立点
- 一阶微分对阶梯状的灰度变化有较强的响应
- 二阶微分在处理阶梯状灰度变化时产生双响应
- 如果灰度的变化相似，二阶微分对线的反应比对阶梯强，对点的反应比对线强

### 3.4.2 拉普拉斯算子

定义 9 (拉普拉斯(Laplacian)算子). 对连续函数情形，最简单且各向同性的二阶微分算子是拉普拉斯算子

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

离散情况

$$\nabla^2 = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)$$

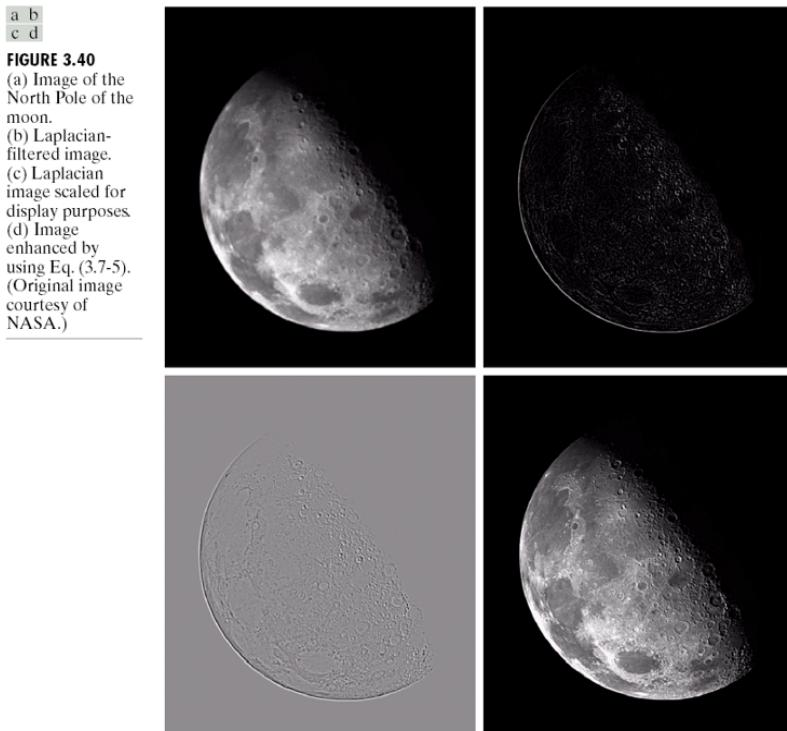
拉普拉斯变换目的是获得细节，加回原图像才能进行锐化。

| a      | b      | c       | d        |
|--------|--------|---------|----------|
| 0 1 0  | 1 -8 1 | 0 -1 0  | -1 -1 -1 |
| 1 -4 1 | 1 1 1  | -1 4 -1 | -1 8 -1  |
| 0 1 0  | 1 1 1  | 0 -1 0  | -1 -1 -1 |

图 3.37 (a) 实现式 (3.6-6) 所用的滤波器模板；(b) 用于实现带有对角项的该公式的扩展的模板；(c)~(d) 实践中常用的其他两个拉普拉斯实现

其实直接从滤波器的表示也可以直观看出这种滤波对图像的突变有比较强的响应（即在突变的位置有较大的输出值），对灰度变化缓慢的区域滤波响应的值会变得很小（变暗）。因此，用拉普拉斯算子作用后，产生的图像将是在暗背景上的一些灰色边线和一些突变点。若将原始图像叠加到拉普拉斯变换后的图像，既可以保护拉普拉斯锐化处理的效果，同时又能复原背景信息。拉普拉斯图像增强基本方法：

$$g(x, y) = \begin{cases} f(x, y) - \nabla^2 f & \text{拉普拉斯滤波中心系数为负} \\ f(x, y) + \nabla^2 f & \text{拉普拉斯滤波中心系数为正} \end{cases}$$



**定义 10** (反锐化掩膜(unsharp masking))。把原图的一个模糊过的图像从原图中减去，从而得到一个相对清晰的图像

$$f_s(x, y) = f(x, y) - \bar{f}(x, y)$$

就像一个模糊的负片和一个正片放在一起冲洗出相对清晰的照片

**定义 11** (高提升滤波(high-boost filtering)). 添加一个系数项  $A$  得到

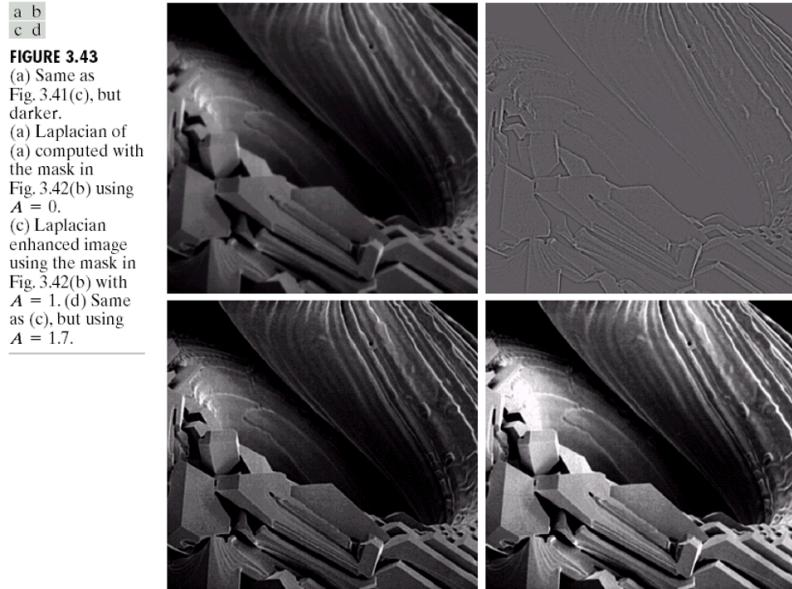
$$f_{hb}(x, y) = Af(x, y) - \bar{f}(x, y) = (A - 1)f(x, y) + f_s(x, y)$$

其中  $A \geq 1$ , 目的为提升原图亮度, 前一部分调整了原图的灰度, 后一部分是锐化过的图像。  $A$  越大则细节越不清晰, 因为原图变亮了。

可以得到基于拉普拉斯算子的高提升滤波

$$f_{hb}(x, y) = \begin{cases} Af(x, y) - \nabla^2 f & \text{Laplacian 中心系数为负} \\ Af(x, y) + \nabla^2 f & \text{Laplacian 中心系数为正} \end{cases}$$

当  $A = 1$  时就是拉普拉斯图像增强方法, 当  $A$  足够大时, 锐化效果将变得不明显。



### 3.4.3 梯度

一阶微分在灰度的跳跃性间断处(边界处)有较强的响应, 所以在一些情况下也可以用于图像增强, 常用作边缘检测。考虑二维函数的梯度

$$\nabla f = [G_x \quad G_y] = \left[ \frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \right]$$

定义  $L_2$  范数/模

$$\|\nabla f\| = (G_x^2 + G_y^2)^{1/2} = \left( \left( \frac{\partial f}{\partial x} \right)^2 + \left( \frac{\partial f}{\partial y} \right)^2 \right)^{1/2}$$

$L_2$  模具有各向同性的性质, 但计算不方便, 故常用  $L_1$  范数做代替

$$\|\nabla f\| \approx |G_x| + |G_y|$$

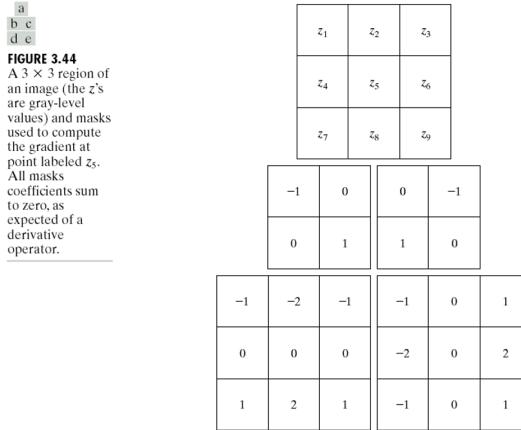
注意：通常在不引起混淆的情况下，把梯度的模称为梯度。

- Robert交叉梯度算子

$$\nabla f = |z_9 - z_5| + |z_8 - z_6|$$

- Sobel算子：水平边缘增强

$$\nabla f = |(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)| + |(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)|$$



通常实际使用时是多种滤波器混合使用。下列为Matlab中预定义的滤波器。

- Gaussian 低通滤波器
- Sobel 水平边缘增强滤波器
- Prewitt 水平边缘增强滤波器：将Sobel的系数绝对值全改为1
- Laplacian 近似二维拉普拉斯运算滤波器
- Log (Laplacian of Gaussian)高斯拉普拉斯滤波器
- Average 均值滤波器
- Unsharp 模糊对比增强滤波器

## 4 频率域滤波

### 4.1 傅里叶级数与傅里叶变换

**定义 12** (傅里叶级数). 设 $f(t)$ 为以 $T$ 为周期的函数，绝对可积，则 $f(t)$ 可表示为

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{j \frac{2\pi n}{T} t}$$

其中 $j$ 为虚数单位，

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-j \frac{2\pi n}{T} t} dt, n = 0, \pm 1, \pm 2, \dots$$

傅里叶级数中每一个基函数都是一个单频谐波，对应的系数/频谱表明原函数对这种频率成分贡献的大小（原函数在这个谐波上的投影）

**定义 13** (冲激/狄利克 $\delta$ 函数). 面积为 1 的长方形不断压扁，最后变成宽为 0，高为无穷大的函数（左侧为连续情况，右侧为离散情况）

$$\delta(t) = \begin{cases} \infty & t = 0 \\ 0 & t \neq 0 \end{cases} \quad \delta(x) = \begin{cases} 1 & x = 0 \\ 0 & x \neq 0 \end{cases}$$

即

$$\int_{-\infty}^{\infty} \delta(t) dx = 1 \quad \sum_{x=-\infty}^{\infty} \delta(x) = 1$$

具有取样(*sifting*)特性

$$\int_{-\infty}^{\infty} f(t) \delta(t - t_0) dx = f(t_0) \quad \sum_{x=-\infty}^{\infty} f(x) \delta(x - x_0) = f(x_0)$$

**定义 14** (冲激串). 无限多个分离的周期冲激单元 $\Delta T$ 之和

$$s_{\Delta T}(t) = \sum_{n=-\infty}^{\infty} \delta(t - n\Delta T)$$

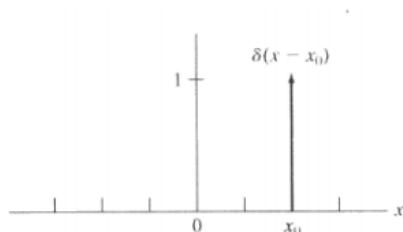


图 4.2 位于  $x = x_0$  处的单位离散冲激。变量  $x$  是离散的，任何不在  $x = x_0$  处的  $\delta$  都是 0

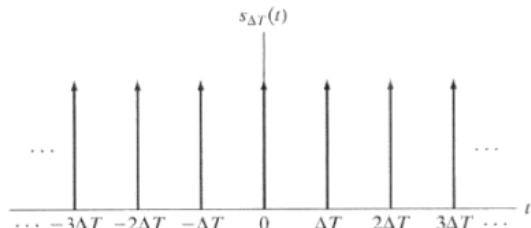


图 4.3 一个冲激串

**定义 15** (傅里叶变换与反变换). 连续情形的傅里叶变换（函数投影，取负号得共轭）和反变换为

$$F(u) = \Im[f(t)] = \int_{-\infty}^{+\infty} f(t) e^{-j2\pi ut} dt$$

$$f(t) = \Im^{-1}[F(u)] = \int_{-\infty}^{+\infty} F(u) e^{j2\pi ut} du$$

这两者构成一个傅里叶变换对。对于离散情形有

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

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

由于傅里叶变换是  $f(t)$  乘上正弦项的展开，正弦项的频率由  $\mu$  决定（变量  $t$  已经被积分），积分后只剩下频率，故称傅里叶变换域是频率域。注意坐标轴已经变化了，现在横轴为频率。

### 例 1. 矩形函数

$$f(t) = \begin{cases} A & -W/2 \leq t \leq W/2 \\ 0 & otherwise \end{cases}$$

的傅里叶变换为

$$\begin{aligned} F(u) &= \int_{-W/2}^{W/2} A e^{-j2\pi ut} dt \\ &= \frac{-A}{j2\pi\mu} [e^{-j2\pi\mu W} - e^{j2\pi\mu W}] \\ &= AW \frac{\sin(\pi\mu W)}{\pi\mu W} = AW \operatorname{sinc}(\mu W) \end{aligned}$$

### 例 2. 冲激的傅里叶变换（由取样特性）

$$F(\mu) = \int_{-\infty}^{\infty} \delta(t) e^{-j2\pi\mu t} dt = e^{-j2\pi\mu 0} = 1$$

而位于  $t = t_0$  的冲激傅里叶变换为

$$F(\mu) = \int_{-\infty}^{\infty} \delta(t - t_0) e^{-j2\pi\mu t} dt = e^{-j2\pi\mu t_0}$$

周期冲激串的傅里叶变换

$$S(\mu) = \Im[s_{\Delta T}(t)] = \frac{1}{\Delta T} \sum_{n=-\infty}^{\infty} \delta\left(\mu - \frac{n}{\Delta T}\right)$$

**定义 16 (卷积).** 连续情形有

$$f(t) * h(t) = \int_{-\infty}^{+\infty} f(\tau) h(t - \tau) d\tau$$

离散情形有

$$f(x) * h(x) = \sum_{m=0}^{M-1} f(m) h(x - m)$$

**定理 1 (卷积定理).** 建立起空间域和频率域<sup>4</sup>的联系

$$f(t) * h(t) \iff F(\mu)H(\mu) \quad f(t)h(t) \iff F(\mu)*H(\mu)$$

即空间域两个函数卷积的傅里叶变换等于两个函数的傅里叶变换在频率域的乘积

---

<sup>4</sup> $t$ 所在的域称为空间域， $\mu$ 所在的域称为频率域

分析.

$$\begin{aligned}
\Im[f(t) * h(t)] &= \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(\tau)h(t-\tau) d\tau \right] e^{-j2\pi\mu t} dt \\
&= \int_{-\infty}^{\infty} f(\tau) \left[ \int_{-\infty}^{\infty} h(t-\tau)e^{-j2\pi\mu t} dt \right] d\tau \\
&= \int_{-\infty}^{\infty} f(\tau) [H(\mu)e^{-j2\pi\mu\tau}] d\tau \\
&= H(\mu) \int_{-\infty}^{\infty} f(\tau) [e^{-j2\pi\mu\tau}] d\tau \\
&= H(\mu)F(\mu)
\end{aligned}$$

## 4.2 取样函数

例 3. 求取样函数

$$\tilde{f}(t) = f(t)s_{\Delta T}(t) = \sum_{n=-\infty}^{\infty} f(t)\delta(t-n\Delta T)$$

的傅里叶变换

分析. 由卷积定理

$$\begin{aligned}
\tilde{F}(u) &= \Im[\tilde{f}(t)] = \Im[f(t)s_{\Delta T}(t)] \\
&= F(u) * S(u) = \int_{-\infty}^{\infty} F(\tau)S(u-\tau) d\tau \\
&= \frac{1}{\Delta T} \int_{-\infty}^{\infty} F(\tau) \sum_{n=-\infty}^{\infty} \delta\left(u-\tau-\frac{n}{\Delta T}\right) d\tau \\
&= \frac{1}{\Delta T} \sum_{n=-\infty}^{\infty} F\left(u-\frac{n}{\Delta T}\right)
\end{aligned}$$

定理 2 (奈奎斯特(Nyquist)采样定理). 如果以超过函数最高频率的两倍的采样率来获得样本，则连续的带限函数<sup>5</sup>可以完全从它的样本集恢复，即

$$\frac{1}{\Delta T} > 2\mu_{\max}$$

若以低于两倍的采样率来采样则会出现混淆现象

---

<sup>5</sup>对于以原点为中心的有限区间（带宽）之外的频率值，其傅里叶变换为零。

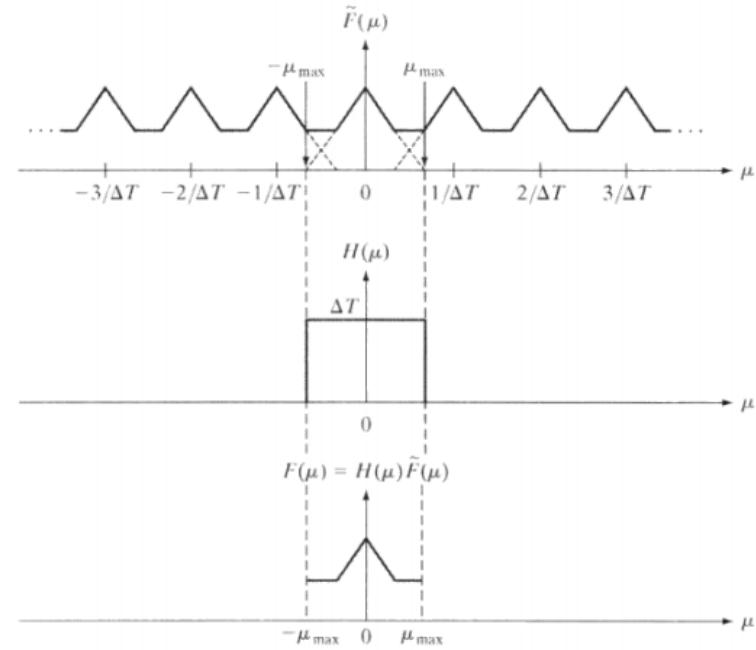


图 4.9 (a)一个欠取样带限函数的傅里叶变换(图中来自邻近周期的干扰显示为虚线); (b)图 4.8(b)中所用的同一个理想低通滤波器; (c)图(a)和图(b)的乘积。来自邻近周期的干扰导致了混淆,而混淆妨碍了  $F(\mu)$  的完美复原,并因此妨碍了原始带限连续函数的完美复原。请与图 4.8 比较

由取样数据重建函数

$$f(t) = \sum_{n=-\infty}^{\infty} f(n\Delta T) \operatorname{sinc} \left[ \frac{t - n\Delta T}{\Delta T} \right]$$

### 4.3 二维傅里叶变换

**定义 17** (二维冲激).

$$\delta(t, z) = \begin{cases} \infty & t = z = 0 \\ 0 & otherwise \end{cases} \quad \delta(x, y) = \begin{cases} 1 & x = y = 0 \\ 0 & otherwise \end{cases}$$

有取样特性

$$\sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} f(x, y) \delta(x - x_0, y - y_0) = f(x_0, y_0)$$

**定义 18** (二维离散傅里叶变换). 二维连续傅里叶变换对

$$\begin{aligned} F(u, v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t, z) e^{-j2\pi(ut+vz)} dt dz \\ f(t, z) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{j2\pi(ut+vz)} du dv \end{aligned}$$

二维离散傅里叶变换对

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-2j\pi(ux/M+vy/N)}$$

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{2\pi j(ux/M+vy/N)}$$

**定理 3** (二维采样定理). 二维取样基于二维冲激串

$$\tilde{f}(t, z) = f(t, z) s_{\Delta T \Delta Z}(t, z) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(t, z) \delta(t - m\Delta T, z - n\Delta Z)$$

取样率需要满足

$$\frac{1}{\Delta T} > 2\mu_{\max} \quad \frac{1}{\Delta Z} > 2v_{\max}$$

**定义 19** (傅里叶谱和相角). 由于二维DFT通常为复函数，因此用极坐标形式表示

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

其中幅度

$$|F(u, v)| = \sqrt{\operatorname{Re}^2(u, v) + \operatorname{Im}^2(u, v)}$$

称为傅里叶谱，而

$$\varphi(u, v) = \arctan \left[ \frac{\operatorname{Im}(u, v)}{\operatorname{Re}(u, v)} \right]$$

为相角，

$$P(u, v) = |F(u, v)|^2$$

为功率谱

注意 $F(u, v)$ 的大小只是代表某一频率分量的数目，类似于直方图的概念。频率域的坐标轴为 $u, v$ ，因此中心点为 $(0, 0)$ ，

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

为图像的平均灰度。即 $u$ 方向和 $v$ 方向上最低频的位置，零点处两个方向的频率都为零，因此这个量经常也被称为频谱的直流分量(DC)。

**例 4** (陷波滤波器(notch filter)). 使得处理后的图像均值为0，从而使图像的整体灰度降低

$$H(u, v) = \begin{cases} 0 & (u, v) = (M/2, N/2) \\ 1 & others \end{cases}$$

#### 例 4.13 一个简单函数的二维傅里叶谱。

图4.24(a)显示了一幅简单的图像, 图4.24(b)显示了该图像的谱, 其值被标定到区间[0, 255], 并以图像形式显示。空间域和频率域的原点都在左上角。在图4.24(b)中, 两件事情是很明显的。如期望的那样, 变换的原点的周围区域包含了最高值(在图像中表现为最亮)。然而, 注意, 谱的4个角包含有类似的高值。原因是在前一节中讨论的周期性。为中央化该谱, 如式(4.6-8)指出的那样, 在计算DFT之前, 我们简单地用 $(-1)^{x+y}$ 乘以(a)中的图像。图4.24(c)显示了结果, 很明显它更易于观察(注意关于中心点的对称性)。由于直流项支配着谱的值, 在显示的图像中, 其他灰度的动态范围被压缩了。为了给出那些细节, 如3.2.2节描述的那样, 我们执行一次对数变换。图4.24(d)给出了 $(1 + \log|F(u, v)|)$ 的显示, 细节再现的提高是明显的。本章和后续几章中的多数谱均按这种方式标定过。

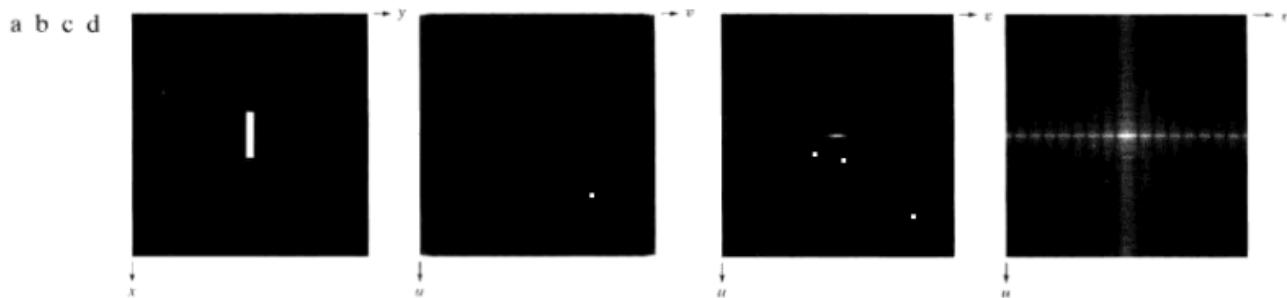


图 4.24 (a) 图像; (b) 在 4 个角显示亮点的谱; (c) 中心化后的谱; (d) 经对数变换后, 增加了细节的结果。谱的零交叉在垂直方向较靠近, 因为图(a)中的矩形在该方向上较长。全书所用的坐标习惯把空间域和频率域的原点放在左上角处

相角对形状起着决定性作用! 灰度信息则由谱携带。

#### 例 4.14 傅里叶谱和相角的进一步说明。

图 4.27(b) 是图 4.27(a) 的 DFT 的相角。该阵列将通过视觉分析引导我们将其与对应图像中的特性关联起来。然而，在决定形状特点时，相位的重要性在图 4.27(c) 中是很明显的，它仅仅用相位信息（也就是，用公式  $|F(u, v)|=1$ ）计算式(4.6-15)的反傅里叶变换得到。尽管灰度信息已丢失了（记住，该信息由谱携带），但图像中关键的形状特性是不会在图 4.27(a) 中弄错的。

图 4.27(d) 是仅使用式(4.6-15)的谱并计算其 IDFT 得到的。这意味着若将指数项置为 1，这也同样意味着置相位角为 0，结果不会是意外的。它仅包含灰度信息，直流项占支配地位。图像中没有形状信息，因为相位已置为零了。

最后，图 4.27(e) 和(f) 再一次证明了在确定一幅图像的特性内容时相位所起的支配作用。图 4.27(e) 是使用图 4.24(a) 中矩形的谱和对应于妇女图像的相角计算式(4.6-15)的 IDFT 得到的。妇女的形状显然支配了这一结果。反之，矩形支配了图 4.27(f)，它是使用妇女图像的谱和矩形的相角计算得到的。

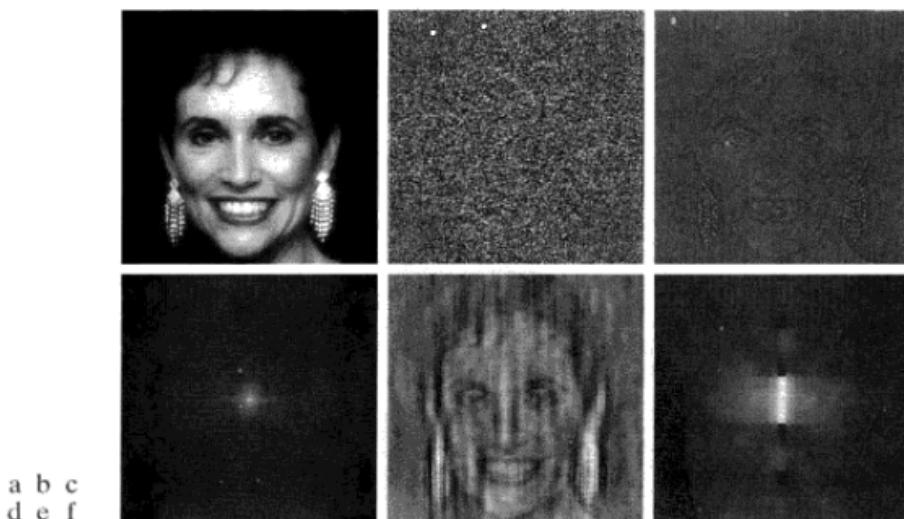


图 4.27 (a) 妇女图像；(b) 相角；(c) 仅使用相角重建的妇女图像；(d) 仅使用谱重建的妇女图像；(e) 使用对应于妇女图像的相角和对应于图 4.24(a) 中矩形的谱重建的妇女图像；(f) 使用矩形的相角和妇女图像的谱重建的图像

**定理 4** (二维卷积定理)。二维的卷积定义如下

$$f(x, y) * h(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n)h(x-m, y-n)$$

同样有傅里叶变换对

$$f(x, y) * h(x, y) \iff F(u, v)H(u, v)$$

$$f(x, y)h(x, y) \iff F(u, v) * H(u, v)$$

注意周期靠近会导致互相干扰而导致缠绕错误，因此需要进行零延拓(padding)。

假设函数 $f(x)$ 和 $h(x)$ 分别有 $A$ 个和 $B$ 个点组成，对这两个函数同时添加零，使其具有相同的周期

$$f_e = \begin{cases} f(x) & 0 \leq x \leq A-1 \\ 0 & A \leq x \leq P \end{cases} \quad h_e = \begin{cases} h(x) & 0 \leq x \leq B-1 \\ 0 & B \leq x \leq P \end{cases}$$

同样对于二维图像滤波，若 $f(x, y)$ 大小为 $A \times B$ ,  $h(x, y)$ 大小为 $C \times D$ , 则延拓函数

$$f_e(x, y) = \begin{cases} f(x, y) & 0 \leq x \leq A-1 \wedge 0 \leq y \leq B-1 \\ 0 & A \leq x \leq P \vee B \leq y \leq Q \end{cases} \quad h_e(x, y) = \begin{cases} h(x, y) & 0 \leq x \leq C-1 \wedge 0 \leq y \leq D-1 \\ 0 & C \leq x \leq P \vee D \leq y \leq Q \end{cases}$$

其中 $P \geq A + C - 1$ ,  $Q \geq B + D - 1$ , 通常可取 $P = 2M$ ,  $Q = 2N$ 。

**定义 20** (相关). 相关性相当于算内积

$$g(x, y) = f(x, y) \circ h(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f^*(m, n) h(x+m, y+n)$$

**定理 5** (相关定理).

$$\begin{aligned} f(x, y) \circ h(x, y) &\iff F^*(u, v) H(u, v) \\ f^*(x, y) h(x, y) &\iff F(u, v) \circ H(u, v) \end{aligned}$$

#### 4.4 频率域滤波

之所以要到频率域做滤波，是因为直观且计算比空间滤波容易（卷积耗时）

- 低通滤波器：平滑/模糊图像，设 $D(u, v)$ 为频率域中点 $(u, v)$ 与频率矩形中心的距离， $D_0$ 为截止频率

| 理想(ILPF)  | $n$ 阶布特沃斯(BLPF)                              | 高斯                                |
|---|--|-----------------------------------|
| $H(u, v) = \begin{cases} 1 & D(u, v) \leq D_0 \\ 0 & D(u, v) > D_0 \end{cases}$ | $H(u, v) = \frac{1}{1 + [D(u, v)/D_0]^{2n}}$ | $H(u, v) = e^{-D^2(u, v)/2D_0^2}$ |

原点在频率中心，半径为 $r$ 的圆包含的功率为

$$\alpha = \left( \sum_u \sum_v \frac{P(u, v)}{P_T} \right) \times 100\% = \left( \sum_u \sum_v \frac{P(u, v)}{\sum_{u=0}^{M-1} \sum_{v=0}^{N-1} |F(u, v)|^2} \right) \times 100\%$$

由于截止频率点处跳变太直接，物理无法实现，故是理想滤波器；而且会产生滤波模糊和振铃现象。一阶BLPF没有振铃，二阶BLPF振铃很小，随着阶数增高振铃增大，故通常用二阶。

- 高通滤波器： $H_{hp}(u, v) = 1 - H_{lp}(u, v)$ ，锐化图像（噪声、边缘、细节），只获得高频特征，没有背景，留下细节

| 理想(IGPF)  | $n$ 阶布特沃斯(BGPF)                              | 高斯                                    |
|---|--|---------------------------------------|
| $H(u, v) = \begin{cases} 0 & D(u, v) \leq D_0 \\ 1 & D(u, v) > D_0 \end{cases}$ | $H(u, v) = \frac{1}{1 + [D_0/D(u, v)]^{2n}}$ | $H(u, v) = 1 - e^{-D^2(u, v)/2D_0^2}$ |

- 拉普拉斯滤波

$$H(u, v) = -4\pi^2 \left[ (u - P/2)^2 + (v - Q/2)^2 \right] = -4\pi^2 D^2(u, v)$$

前面的系数可以去掉。拉普拉斯图像由下式给出

$$\nabla^2 f(x, y) = \mathfrak{I}^{-1}[H(u, v)F(u, v)]$$

图像增强可以由下式实现

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

- 高频增强滤波：锐化/加强图像，高频增强，细节变得明显

$$\begin{aligned} F_{lp}(u, v) &= H_{lp}(u, v)F(u, v) \\ F_{hp}(u, v) &= F(u, v) - F_{lp}(u, v) \\ &= (1 - H_{lp}(u, v))F(u, v) \\ &= H_{hp}(u, v)F(u, v) \\ G(u, v) &= F(u, v) + F_{hp}(u, v) \\ &= (1 + H_{hp}(u, v))F(u, v) \end{aligned}$$

其中  $k = 1$  为钝化模板， $k > 1$  为高提升滤波

- 同态滤波器：高频增强，整个图像亮度又不能太亮（一片的亮度与低频有关） $\implies$  抑制低频/环境光，提升高频/细节

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

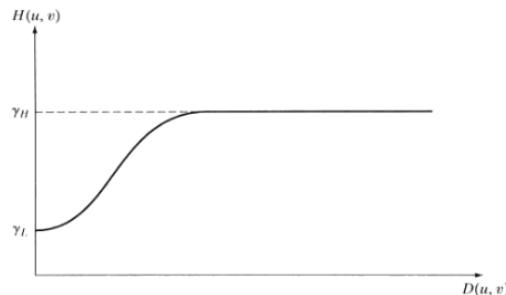
其中  $i(x, y) \in (0, \infty)$  为照射分量（影响低频，一片）； $0 < r(x, y) < 1$ ，反射率影响边缘/高频。先取对数，然后再做傅里叶变换，最后记得取指数返回

$$\ln f(x, y) = \ln i(x, y) + \ln r(x, y)$$

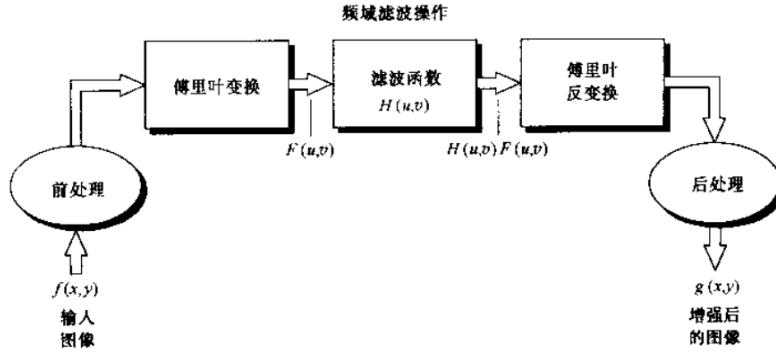
同态滤波器

$$H(u, v) = (r_H - r_L)[1 - e^{-c(D^2(u, v)/D_0^2)}] + r_L$$

其中  $c$  用于控制滤波器函数斜面的锐化， $r_L < 1$ ， $r_H > 1$



## 频率域滤波的步骤



1. 给定大小为  $M \times N$  的输入图像  $f(x, y)$ , 选择填充参数  $P = 2M, Q = 2N$
2. 对  $f(x, y)$  添加必要的0, 形成大小为  $P \times Q$  的填充后的图像  $f_p(x, y)$
3. 用  $(-1)^{x+y}$  乘以  $f_p(x, y)$ , 做频谱中心化处理
4. 用上面结果计算DFT, 得到  $F(u, v)$
5. 生成一个实对称的滤波函数  $H(u, v)$ , 大小为  $P \times Q$ , 中心在  $(P/2, Q/2)$  处。用阵列相乘得到  $G(u, v) = H(u, v)F(u, v)$
6. 计算上式得到的IDFT, 并恢复原图像

$$g_p(x, y) = \operatorname{Re}[\Im[G(u, v)]](-1)^{x+y}$$

7. 通过从  $g_p(x, y)$  的左上角提取  $M \times N$  大小的区域, 得到最终结果  $g(x, y)$

## 4.5 总结

傅里叶变换的一些性质

- 线性性:

$$af_1(x, y) + bf_2(x, y) \iff aF_1(u, v) + bF_2(u, v)$$

- 平移性质:

$$\begin{aligned} f(x, y)e^{j2\pi(u_0x/M+v_0y/N)} &\iff F(u - u_0, v - v_0) \\ f(x - x_0, y - y_0) &\iff F(u, v)e^{-2j\pi(x_0u/M+y_0v/N)} \end{aligned}$$

特别地, 平移到频率矩形中心  $(M/2, N/2)$

$$\begin{aligned} f(x, y)(-1)^{x+y} &\iff F(u - M/2, v - N/2) \\ f(x - M/2, y - N/2) &\iff F(u, v)(-1)^{u+v} \end{aligned}$$

- 旋转性质: 使用极坐标  $x = r \cos \theta, y = r \sin \theta, u = \omega \cos \varphi, v = \omega \sin \varphi$ , 有

$$f(r, \theta + \theta_0) \iff F(\omega, \varphi + \varphi_0)$$

- 周期性:

$$F(u, v) = F(u + k_1 M, v + k_2 N)$$

$$f(x, y) = f(x + k_1 M, y + k_2 N)$$

- 对称性: 实函数  $f(x, y)$  的傅里叶变换是共轭对称/哈密特对称的

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

而虚函数  $f(x, y)$  的傅里叶变换是共轭反对称的

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

- 可分性: 2D-FFT 可变成两个1D-FFT

$$F(u, v) = \sum_{x=0}^{M-1} \left( \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi ux/M} \right) e^{-j2\pi vy/N}$$

常见函数的傅里叶变换

|        |  |
|--------|--|
| 离散单位冲激 | $\delta(x, y) \iff 1, 1 \iff \delta(u, v)$   |
| 矩形函数   | $\text{rect}[a, b] \iff ab \frac{\sin(\pi ua)}{\pi ua} \frac{\sin(\pi ub)}{\pi ub} e^{-j\pi(ua+vb)}$   |
| 正弦函数   | $\sin(2\pi u_0 x + 2\pi v_0 y) \iff j\frac{1}{2}[\delta(u + Mu_0, v + Nv_0) - \delta(u - Mu_0, v - Nv_0)]$   |
| 余弦函数   | $\cos(2\pi u_0 x + 2\pi v_0 y) \iff j\frac{1}{2}[\delta(u + Mu_0, v + Nv_0) + \delta(u - Mu_0, v - Nv_0)]$   |
| 微分     | $\frac{d^n f(x)}{dx^n} \iff (ju)^n F(u), \left(\frac{\partial}{\partial t}\right)^m \left(\frac{\partial}{\partial z}\right)^n f(t, z) \iff (j2\pi u)^m (j2\pi v)^n F(u, v)$ |
| 高斯     | $A2\pi\sigma^2 e^{-2\pi^2\sigma^2(t^2+z^2)} \iff Ae^{-(u^2+v^2)/2\sigma^2}$  |

用前向变换计算傅里叶反变换, 以消除两套冗余电路

$$MNf^*(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F^*(u, v) e^{-j2\pi(ux/M+vy/N)}$$

## 5 图像复原

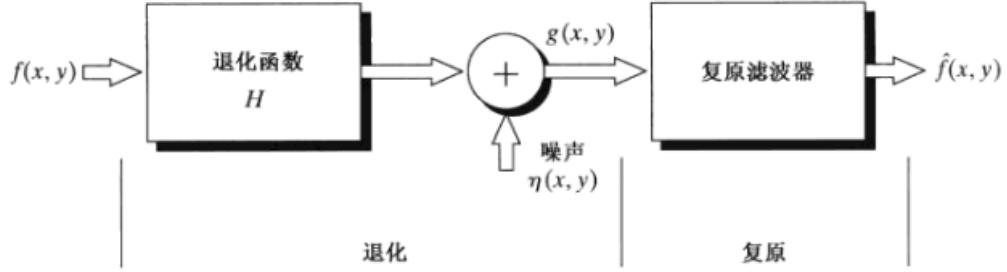
图像增强是主观的, 图像复原则是客观的 (恢复原图)。

空间域和频率域的退化图像分别可以由下式的退化模型给出

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

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

关键在于怎么找回  $f$  和  $F$ , 而  $h$  属于系统噪声。



## 5.1 噪声模型

- 高斯噪声

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\bar{z})^2}{2\sigma^2}}$$

- 脉冲/椒盐噪声

$$p(z) = \begin{cases} P_a & z = a \\ P_b & z = b \\ 1 - P_a - P_b & \text{others} \end{cases}$$

- 瑞利噪声: 均值  $\bar{z} = a + \sqrt{\pi b/4}$ , 方差  $\sigma^2 = b(4 - \pi)/4$

$$p(z) = \begin{cases} \frac{2}{b}(z-a)e^{-(z-a)^2/b} & z \geq a \\ 0 & z < a \end{cases}$$

- 爱尔兰/伽马噪声: 均值  $\bar{z} = b/a$ , 方差  $\sigma^2 = b/a^2$

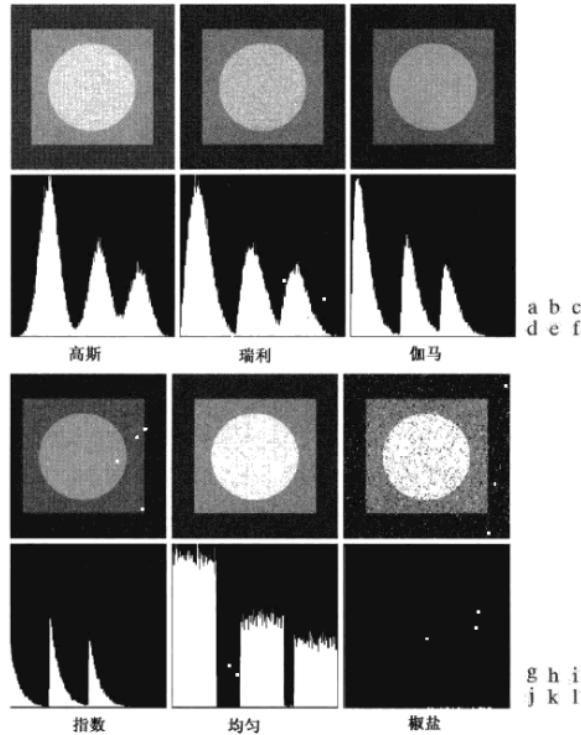
$$p(z) = \begin{cases} \frac{a^b z^{b-1}}{(b-1)!} e^{-az} & z \geq a \\ 0 & z < a \end{cases}, \quad a > 0, b \in \mathbb{Z}^+$$

- 指数噪声: 均值  $\bar{z} = 1/a$ , 方差  $\sigma^2 = 1/a^2$

$$p(z) = \begin{cases} ae^{-az} & z \geq 0 \\ 0 & z < 0 \end{cases}$$

- 均匀噪声: 均值  $\bar{z} = (a+b)/2$ , 方差  $\sigma^2 = (b-a)^2/12$

$$p(z) = \begin{cases} \frac{1}{b-a} & a \leq z \leq b \\ 0 & \text{others} \end{cases}$$



## 5.2 噪声存在下的空间滤波复原

考虑唯一存在噪声退化，即不存在系统退化，则原式变为

$$g(x, y) = f(x, y) + \eta(x, y)$$

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

### 5.2.1 均值滤波器

- 算术均值滤波

$$\hat{f}(x, y) = \frac{1}{MN} \sum_{(s,t) \in S_{xy}} g(s, t)$$

- 几何均值滤波

$$\hat{f}(x, y) = \left[ \prod_{(s,t) \in S_{xy}} g(s, t) \right]^{\frac{1}{mn}}$$

相比算术均值，丢失图像细节比较少

- 谐波均值滤波

$$\hat{f}(x, y) = \frac{mn}{\sum_{(s,t) \in S_{xy}} \frac{1}{g(s, t)}}$$

对盐(大的)噪声比较好，对胡椒噪声不好，善于处理高斯噪声

- 逆谐波均值滤波

$$\hat{f}(x, y) = \frac{\sum_{(s,t) \in S_{xy}} g(s, t)^{Q+1}}{\sum_{(s,t) \in S_{xy}} g(s, t)^Q}$$

其中 $Q$ 称为滤波器的阶数，适合减少椒盐噪声的影响。当 $Q$ 为正时，消除胡椒噪声；当 $Q$ 为负时，消除盐粒噪声；但不能同时消除这两种噪声。当 $Q = 0$ 时退化为算术均值， $Q = -1$ 为谐波均值滤波

### 5.2.2 状态/顺序统计滤波器

- 最大滤波器
- 最小滤波器
- 中点滤波器：

$$\hat{f}(x, y) = \frac{1}{2} \left[ \max_{(x,y) \in S_{xy}} g(s, t) + \min_{(x,y) \in S_{xy}} g(s, t) \right]$$

- 修正阿尔法均值滤波器：去除 $d/2$ 个最大值（去盐）和 $d/2$ 最小值（去椒），再取平均（去高斯）

$$\hat{f}(x, y) = \frac{1}{mn - d} \sum_{(x,y) \in S_{xy}} g_r(s, t)$$

### 5.2.3 自适应滤波器

滤波器作用于局部区域 $S_{xy}$ ， $\sigma_\eta^2$ 为干扰 $f(x, y)$ 以形成 $g(x, y)$ 的噪声方差， $m_L$ 为局部均值， $\sigma_L^2$ 为局部方差

$$\hat{f}(x, y) = g(x, y) - \frac{\sigma_\eta^2}{\sigma_L^2} [g(x, y) - m_L]$$

满足以下假设

- 如果 $\sigma_\eta^2$ 为0，则应直接返回 $g(x, y)$ 的值
- 如果局部方差与 $\sigma_\eta^2$ 高度相关，则滤波器返回 $g(x, y)$ 的近似值。典型地，高局部方差与边缘相关，应该保护这些边缘
- 如果两个方差相等，则希望滤波器返回 $S_{xy}$ 中像素的算术均值。这种情况发生在局部区域与整个图像有相同特性条件下，局部噪声将通过简单地求平均来降低。

## 5.3 通过频域滤波抑制噪声

### 5.3.1 带阻滤波器

只有中间 $W$ 宽度的能够通过，同样有理想带阻滤波器、布特沃斯滤波器、高斯带阻滤波器

$$H(u, v) = \begin{cases} 1 & D(u, v) < D_0 - \frac{W}{2} \vee D(u, v) > D_0 + \frac{W}{2} \\ 0 & \text{otherwise} \end{cases}$$

带阻滤波可以消除周期性噪声。

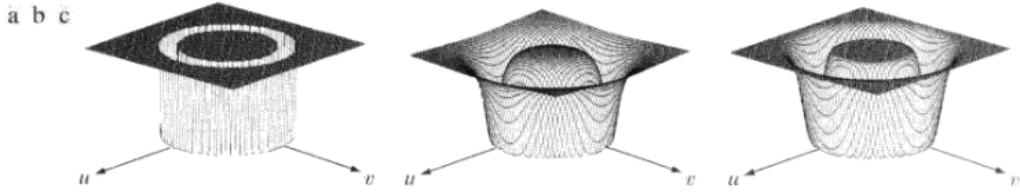


图 5.15 从左到右依次为理想滤波器、布特沃斯滤波器(阶数为 1)和高斯带阻滤波器的透视图

### 5.3.2 带通滤波器

$$H_{bp}(u, v) = 1 - H_{br}(u, v)$$

只是将带阻滤波器取反

### 5.3.3 陷波滤波器

阻止(或通过)事先定义得中心频率邻域内的频率

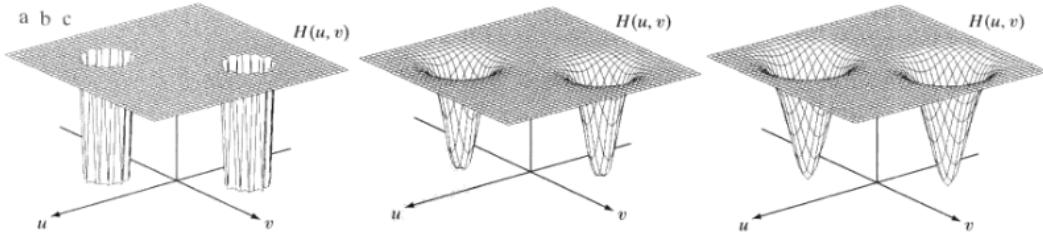


图 5.18 (a) 理想、(b) 布特沃斯(2 阶) 和(c) 高斯陷波(带阻) 滤波器的透视图

由于频率域的共轭对称性，去掉一个，要去除另一个对称的频率。

陷波带通滤波器可以用于确定噪声模式。

如何评判图像的好坏

- 方差最小，噪声水平最低
- 是不是平滑的

## 5.4 线性、位置不变的退化

输入/输出关系可以表示为

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

具有加性噪声的线性空间不变退化系统，可在空间域建模为退化(点扩散)函数与一幅图像的卷积，然后再加上噪声。基于卷积定理，在频率域中，同样的过程可表示为图像和退化函数的变换的乘积，然后再加上噪声的变换。

## 5.5 估计退化函数

- 观察法

$$H_s(u, v) = \frac{G_s(u, v)}{\hat{F}_s(u, v)}$$

- 实验法：获取与退化图像类似的装置，由成像一个亮点得到退化的冲激响应

$$A\delta(x, y) \rightarrow g(x, y) \iff A \rightarrow G(u, v)$$

故  $H(u, v) = \frac{G(u, v)}{A}$

- 数学建模法：如运动模糊

$$g(x, y) = \int_0^T f[x - x_0(t), y - y_0(t)] dt$$

最终得到退化函数

$$H(u, v) = \int_0^T e^{-j2\pi[u x_0(t) + v y_0(t)]} dt$$

## 5.6 逆滤波

对于线性移不变系统的退化模型为

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

如果知道  $H(u, v)$ ，则可以计算原始图像的傅里叶变换估计（去卷积）

$$\hat{F}(u, v) = \frac{G(u, v)}{H(u, v)} = F(u, v) + \frac{N(u, v)}{H(u, v)}$$

如果  $H(u, v)$  很小，则  $N(u, v)/H(u, v)$  将很大，估计就会失败。

维纳滤波：最佳去噪滤波器

$$\hat{F}(u, v) = \frac{P_f(u, v)}{P_f(u, v) + P_k(u, v)} = \frac{P_f(u, v)}{P_f(u, v) + \left| \frac{N(u, v)}{H(u, v)} \right|}$$

## 6 彩色图像处理

图像主要包括三方面内容：颜色、形状、纹理

### 6.1 颜色特性

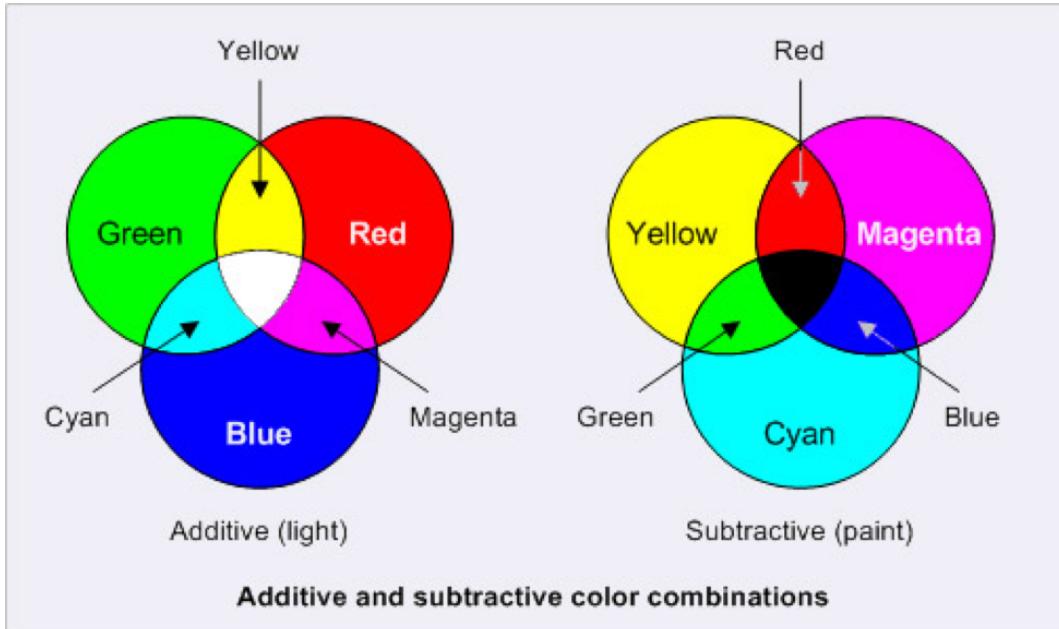
彩色光的3个基本量：

- 辐射率：从光源流出能量的总量，用Watt表示

- 光强：观察者从光源接受的能量总和，用流明表示
- 亮度：主观描绘子，人感觉到的

人眼中约有600-700万个锥状体分别对红色(700nm)、绿色(546nm)、蓝色(435.8nm)敏感<sup>6</sup>。65%对红光敏感、33%对绿光敏感、2%对蓝光敏感。

三基色原理：红绿蓝为三种基色，组成RGB三维加性空间



区别颜色的特性：亮度、色调、色饱和度。颜色通常用亮度和彩度表征，色调、和饱和度统称为彩度。设 $X, Y, Z$ 为红绿蓝的系数，则做归一化，得到颜色的唯一比例

$$x = \frac{X}{X + Y + Z}$$

$$x + y + z = 1$$

真彩色(24b)：RGB各用一个字节表示，共216种安全色（各种设备都可以正常显示），剩下40个为控制字节

## 6.2 颜色空间

常见的颜色空间<sup>7</sup>如下

- RGB：加性空间，与人眼视觉系统密切相连。
  - RGBA：加上透明通道
  - sRGB(standard)：Microsoft
  - Adobe RGB：打印出来更接近原色

<sup>6</sup>注意这里只是人为定义波长，真正的红绿蓝是一段波长区间。

<sup>7</sup>[https://en.wikipedia.org/wiki/List\\_of\\_color\\_spaces\\_and\\_their\\_uses](https://en.wikipedia.org/wiki/List_of_color_spaces_and_their_uses)

- CMY/CMYK：青色(cyan)、品红(magenta)、黄色(yellow)、黑色(key/black)，主要用于打印。打印主要靠反射（**减性空间**），如黄色是白光将蓝色吸收掉。由于油墨很少能将颜色都吸收掉，深色效果较差，故加入一种黑色K。
- HIS/HSL/HSV：色度(hue)、亮度(intensity)、饱和度(saturation)，亮度与色彩分离。广泛应用于计算机视觉、图像检索和视频检索。
- CIE：第一个基于人类视觉感知的颜色空间。CIE-XYZ(1931)是在RGB系统的基础上，用数学方法，选用三个理想的原色来代替实际的三原色，从而将CIE-RGB系统中的光谱三刺激值和色度坐标r、g、b均变为正值。
- Lab：Lab模式既不依赖光线，也不依赖于颜料，它是CIE组织确定的一个理论上包括了人眼可以看见的所有色彩的色彩模式。Lab模式弥补了RGB和CMYK两种色彩模式的不足。同RGB颜色空间相比，Lab是一种不常用的色彩空间。它是一种设备无关的颜色系统，也是一种基于生理特征的颜色系统。这也就意味着，它是用数字化的方法来描述人的视觉感应。Lab颜色空间中的L分量用于表示像素的亮度，取值范围是[0, 100]，表示从纯黑到纯白；a表示从红色到绿色的范围，取值范围是[127, -128]；b表示从黄色到蓝色的范围，取值范围是[127, -128]。
- YUV：明亮度(Y, Luminance/Luma)即灰阶值、色度(U/V, Chrominance/Chroma)用于指定像素颜色。包括YCbCr、YPbPr、YUV、Y'UV等，后两者通常用来编码电视的模拟信号<sup>8</sup>，YCbCr用来描述数字的视频信号，适合视频与图片压缩以及传输，例如MPEG、JPEG。

关于不同颜色空间的优缺点可见<sup>9</sup>。

### 6.3 伪彩色处理

根据一定的准则对灰度值赋以彩色的处理。之所以需要伪彩色，是因为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。比如将不同灰度赋予不同颜色，得到热图(heatmap)。

### 6.4 全彩色处理

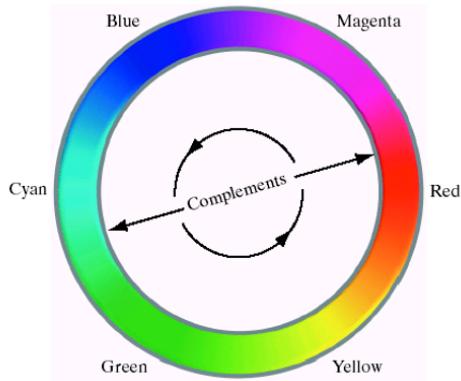
两种处理方式：分通道处理、向量处理。

补色：两种颜色混在一起为白色（RGB补色为CMY）

---

<sup>8</sup>Y'UV的发明是由于彩色电视与黑白电视的过渡时期。黑白视频只有Y(Luma/Luminance)视频，也就是灰阶值。到了彩色电视规格的制定，是以YUV/YIQ的格式来处理彩色电视图像，把UV视作表示彩度的C(Chrominance或Chroma)，如果忽略C信号，那么剩下的Y(Luma)信号就跟之前的黑白电视频号相同，这样一来便解决彩色电视机与黑白电视机的兼容问题。Y'UV最大的优点在于只需占用极少的带宽。

<sup>9</sup><https://blog.csdn.net/JiangHui1211/article/details/84592774>



**FIGURE 6.32**  
Complements on  
the color circle.

## 6.5 彩色分割

- HSI颜色空间分割直观，H色调图像方便描述彩色，S饱和度图像做模板分离感兴趣的特征区，I强度图像不携带彩色信息。如用门限产生二值图像，大于门限的像素赋为1，其他赋为0。
- RGB彩色空间直接，用欧式距离度量。

如果直接采用3个独立平面形成的合成梯度图可能导致彩色边缘检测错误，因此要采用Di Zenzo提出的方法。

## 7 形态学图像处理

### 7.1 腐蚀与膨胀

**定义 21** (反射).

$$\hat{B} = \{w \mid w = -b, b \in B\}$$

**定义 22** (平移).

$$(B)_z = \{c \mid c = b + z, b \in B\}$$

**定义 23** (腐蚀). 结构元 $B$ 可以完全含于 $A$ 中

$$A \ominus B = \{z \mid (B)_z \subset A\} = \{z \mid (B)_z \cap A^C = \emptyset\}$$

注意最终结果只保留结构元的中心点。腐蚀可以消除图像中的某些部分。

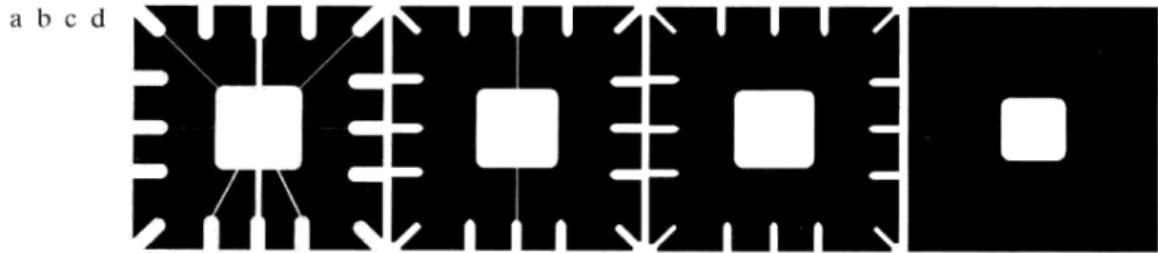


图 9.5 使用腐蚀去除图像中的部件: (a)一幅大小为 486×486 的连线模板二值图像; (b)~(d) 分别使用大小为 11×11, 15×15 和 45×45 的结构元腐蚀的图像。SE 的元素都是 1

**定义 24 (膨胀).** 结构元 $B$ 与 $A$ 有交

$$A \oplus B = \{z \mid (\hat{B})_z \cup A \neq \emptyset\} = \{z \mid [(\hat{B})_z \cap A] \subset A\}$$

膨胀可以用来桥接裂缝。注意要先取反射。

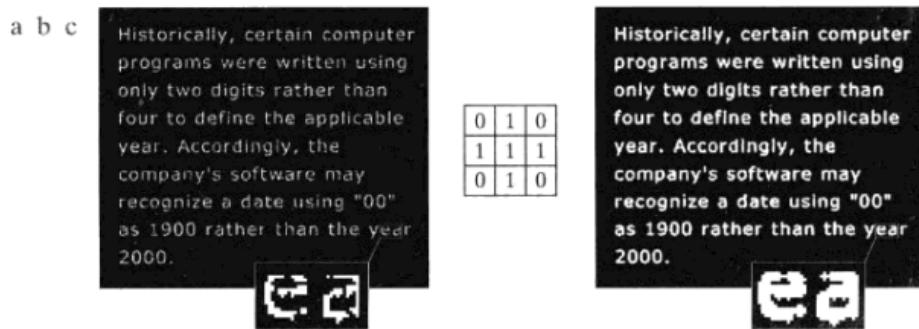


图 9.7 (a) 具有断裂字符的低分辨率样品文本(见放大的视图);  
(b) 结构元; (c) 图(b)对(a)的膨胀。断裂线段被连起来了

膨胀与腐蚀具有对偶性

$$(A \ominus B)^c = A^c \oplus B$$

$$(A \oplus B)^c = A^c \ominus \hat{B}$$

## 7.2 开操作与闭操作

**定义 25 (开操作).**  $B$ 先对 $A$ 腐蚀, 再进行膨胀

$$A \circ B = (A \ominus B) \oplus B = \bigcup \{(B)_z \mid (B)_z \subset A\}$$

实际上就是结构元 $B$ 在 $A$ 内部划过的边界

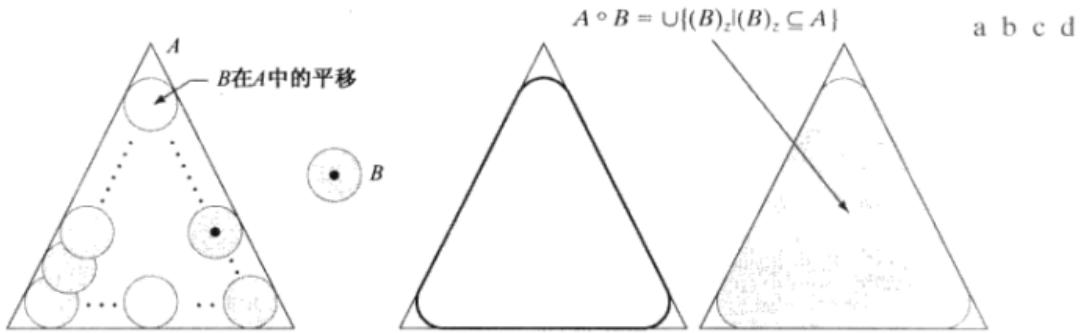


图 9.8 (a) 结构元  $B$  沿集合  $A$  的内侧边界滚动(黑点表示  $B$  的原点); (b) 结构元; (c) 粗线是开操作的外部边界; (d) 完全的开操作(阴影部分)。为清楚起见, 在图(a)中我们未对  $A$  加阴影

**定义 26** (闭操作).  $B$  先对  $A$  膨胀, 在进行腐蚀

$$A \bullet B = (A \oplus B) \ominus B$$

实际上是结构元  $B$  在  $A$  外部划过的边界

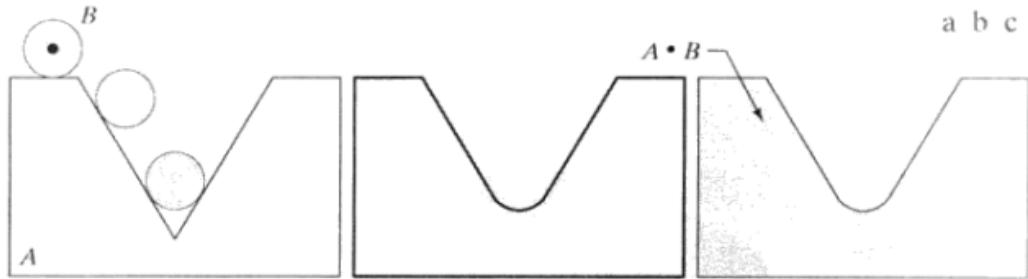


图 9.9 (a) 结构元  $B$  沿集合  $A$  的外侧边界滚动; (b) 粗线是闭操作的外部边界; (c) 完全的闭操作(阴影部分)。为清楚起见, 在图(a)中我们未对  $A$  加阴影

同样, 开操作和闭操作也具有对偶性

$$(A \bullet B)^c = (A^c \circ \hat{B})$$

$$(A \circ B)^c = (A^c \bullet \hat{B})$$

开操作和闭操作满足

- $A \circ B \subset A, A \subset A \bullet B$
- 若  $C \subset D$ , 则  $C \circ B \subset D \circ B, C \bullet B \subset D \bullet B$
- $(A \circ B) \circ B = A \circ B, (A \bullet B) \bullet B = A \bullet B$

### 7.3 击中或击不中变换

**定义 27** (击中或击不中(hit-or-miss)变换).

$$A \circledast B = (A \ominus D) \cap [A^C \ominus (W - D)] = (A \ominus B) \cap (A^C \ominus B_2)$$

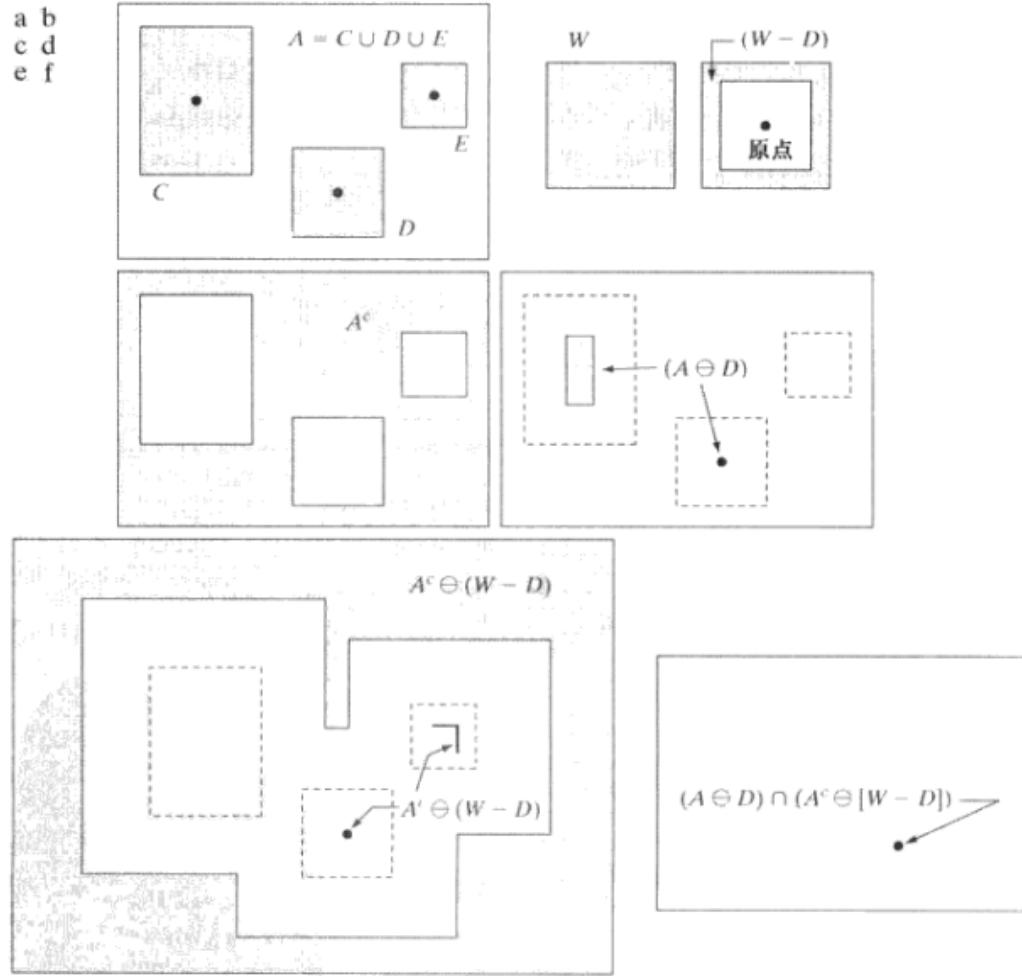


图 9.12 (a)集合  $A$ ; (b)窗口  $W$  和关于  $W$  的  $D$  的局部背景  $(W - D)$ ; (c) $A$  的补集; (d) $D$  对  $A$  的腐蚀; (e) $(W - D)$ 对  $A^c$  的腐蚀; (f) (d) 和 (e) 的交集, 如所希望的那样, 该交集显示了  $D$  的原点的位置。黑点是  $C, D$  和  $E$  的原点

## 7.4 基本形态学算法

- 边界提取

$$\beta(A) = A - (A \ominus B)$$

- 孔洞填充: 当  $X_k = X_{k-1}$  时迭代结束

$$X_k = (X_{k-1} \oplus B) \cap A^C$$

- 连通分量: 当  $X_k = X_{k-1}$  时迭代结束

$$X_k = (X_{k-1} \oplus B) \cap A$$

- 凸包/凸壳: 令 $B^i, i = 1, 2, 3, 4$ 为下列结构元

|       |       |       |       |
|-------|-------|-------|-------|
|       |       |       |       |
| $B^1$ | $B^2$ | $B^3$ | $B^4$ |

$$X_k^i = (X_{k-1} \circledast B^i) \cup A \quad i = 1, 2, 3, 4 \quad k = 1, 2, \dots$$

其中 $X_0^i = A$ , 当该过程收敛 $X_k = X_{k-1}^i$ 时, 令 $D^i = X_k^i$ , 则 $A$ 的凸壳为

$$C(A) = \bigcup_{i=1}^4 D^i$$

凸缺为凸包减去原集合

- 细化

$$A \otimes B = A - (A \circledast B) = A \cap (A \circledast B)^C$$

- 粗化

$$A \odot B = A \cup (A \circledast B)$$

- 骨架

$$S(A) = \bigcup_{k=0}^K S_k(A)$$

其中

$$S_k(A) = (A \ominus kB) - (A \ominus kB) \circ B$$

- 裁剪

$$X_1 = A \otimes \{B\}$$

$$X_2 = \bigcup_{k=1}^8 (X_1 \circledast B^k)$$

$$X_3 = (X_2 \oplus H) \cap A$$

$$X_4 = X_1 \cup X_3$$

## 7.5 灰度级形态学

**定义 28** (灰度级腐蚀与膨胀). 对非平坦结构元 $b_N$ 的腐蚀

$$[f \ominus b](x, y) = \min_{(s,t) \in b} \{f(x+s, y+t)\} - b_N(s, t)$$

对非平坦结构元 $b_N$ 的膨胀

$$[f \oplus b](x, y) = \max_{(s,t) \in b} \{f(x-s, y-t)\} + b_N(s, t)$$

开操作与闭操作都有类似的定义及结论。

**定义 29** (形态学梯度).

$$g = (f \oplus b) - (f \ominus b)$$

**定义 30** (顶帽与底帽变换). 顶帽(*top-hat*)变换

$$T_{hat}(f) = f - (f \circ b)$$

底帽(*bottom-hat*)变换

$$B_{hat}(f) = (f \bullet b) - f$$

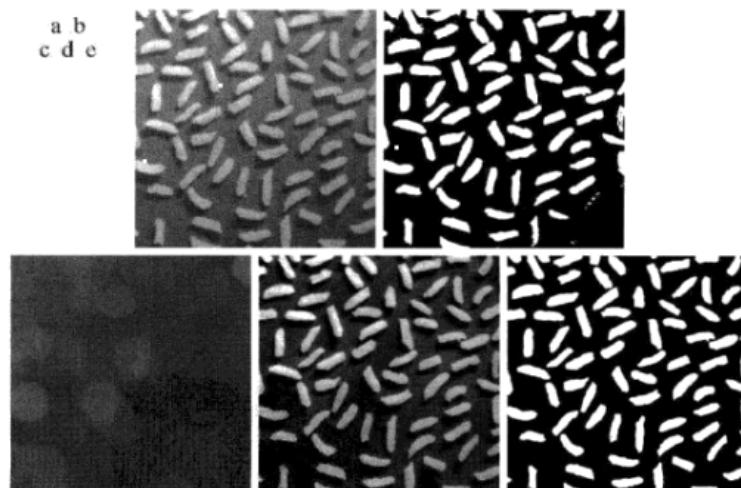


图 9.40 使用顶帽变换校正阴影: (a) 大小为 600x600 的原图像; (b) 阈值处理后的图像; (c) 用半径为 40 的圆盘形结构元进行开操作的图像; (d) 顶帽变换的图像(图像减去其开操作); (e) 顶帽变换的图像经阈值处理的结果

粒度测定原理: 以某一特定的尺度对含有相近尺度颗粒的图像区域进行开操作, 然后通过计算输入图像和输出图像之间的差异可以对相近尺寸颗粒的相对数量进行测算。

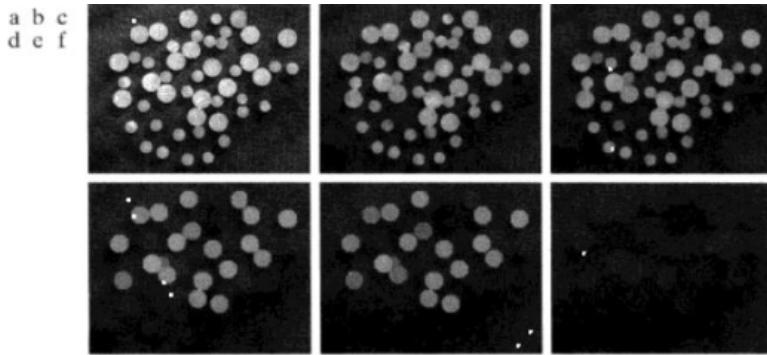


图 9.41 (a) 大小为  $531 \times 675$  的木钉图像; (b) 平滑后的图像; (c)~(f) 分别用半径为 10, 20, 25 和 30 像素的圆盘形结构元对图像进行开操作后的结果 (原图像由 MathWorks 公司的 Steve Eddins 博士提供)

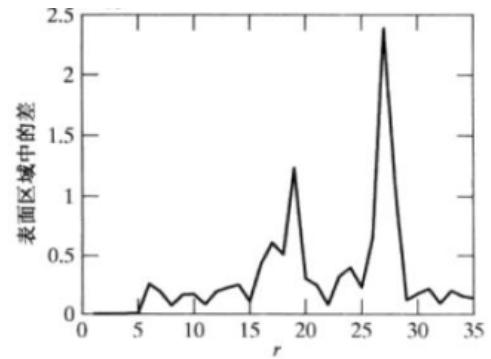


图 9.42 表面区域中的差, 它是圆盘形结构元半径  $r$  的函数。两个峰值表明图像中存在两个主要的颗粒尺寸

纹理分割: 右边区域的圆点直径比左边大, 目的是以纹理为基础找到区域的边界。算法如下:

- 取尺寸与小斑点大小的结构元素做闭运算 (半径30)
- 取比大斑点间隙大的结构元素做开操作 (半径60)
- 做二值化 (使用全1的 $3 \times 3$ 结构元执行梯度画界)



图 9.43 纹理分割: (a) 由两种斑点组成的大小为  $600 \times 600$  的图像; (b) 对图(a)执行闭操作后删除了小斑点的图像; (c) 对图(b)执行开操作后删除了大斑点间的亮间隔的图像; (d) 将图(c)中两个区域间的边界叠加到原图像上后的结果。边界是使用形态学梯度操作得到的

## 8 图像分割

图像分割一般基于亮度值的两种基本特性 (不连续性和相似性) 来分割。

**定义 31 (分割).** 令  $R$  为全图, 可将分割看作  $R$  划分为  $n$  个子区域  $R_1, R_2, \dots, R_n$  的过程:

- $\bigcup_{i=1}^n R_i = R$
- $R_i$  是一个连通区域
- $R_i \cap R_j = \emptyset$
- $Q(R_i) = \text{TRUE}$
- $Q(R_i \cup R_j) = \text{FALSE}$ , 对于任何  $R_i$  和  $R_j$  的邻接区域

其中,  $Q(R_k)$  为定义在集合  $R_k$  的点上的逻辑属性

## 8.1 点检测与线检测

### 8.1.1 点检测

|   |    |   |
|---|----|---|
| 1 | 1  | 1 |
| 1 | -8 | 1 |
| 1 | 1  | 1 |

若作用算子后的图像 $|R(x, y)| \geq T$ , 则记为1。

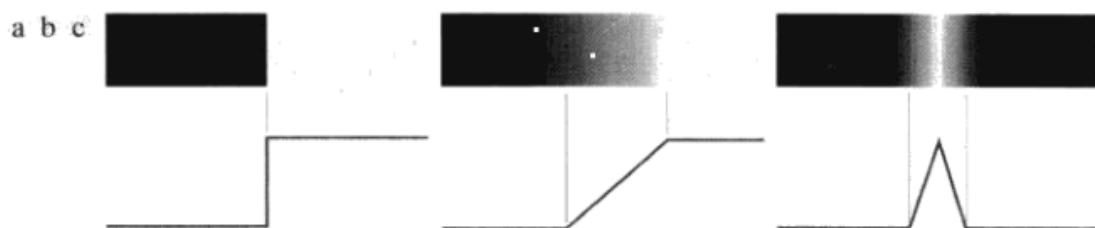
### 8.1.2 线检测

|  |  |  |  |
|--|--|--|--|
| $\begin{array}{ccc} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{array}$ | $\begin{array}{ccc} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{array}$ | $\begin{array}{ccc} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{array}$ | $\begin{array}{ccc} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{array}$ |
| 水平   | $+45^\circ$  | 垂直   | $-45^\circ$  |

## 8.2 边缘检测

傅里叶变换无法刻画边缘, 只知道高频成分, 不知道高频在哪里。一种方法是局部傅里叶变换, 衍生出小波变换(就是要构造一种高通滤波器): 有震荡信号的位置(小范围震荡且积分为0), 可以刻画边缘。

台阶/阶梯、斜坡/斜、屋顶/Delta边缘模型如下。



如下图所示, 二阶导数会增大噪声, 因此做边缘检测之前应该先抑制噪声(平滑)。

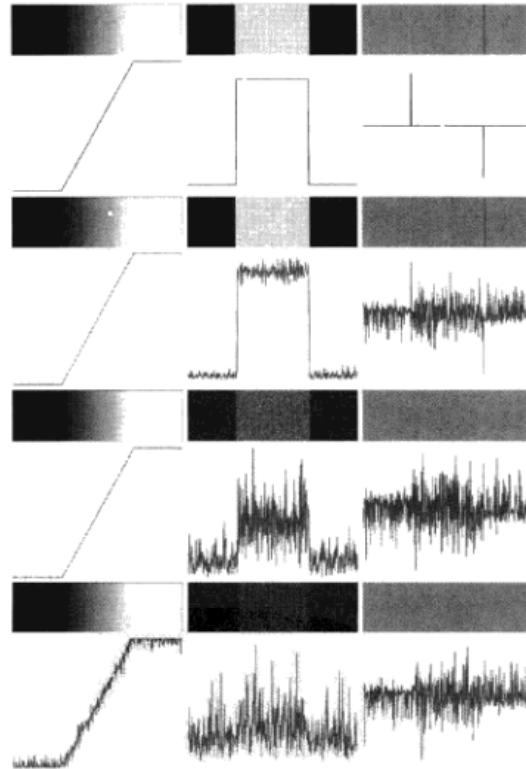


图 10.11 第一列：被均值为零、标准差分别为 0.0, 0.1, 1.0 和 10.0 个灰度级的随机高斯噪声所污染的斜坡边缘的图像和灰度剖面。第二列：一阶导数图像和灰度剖面线。第三列：二阶导数图像和灰度剖面线

常用的边缘检测算子：一阶微分-梯度，二阶微分-Laplace 算子

|  |   |  |  |
|--|---|--|--|
| $\begin{matrix} a & d & e \\ b & c & f & g \end{matrix}$ | $\begin{array}{ c c c } \hline z_1 & z_2 & z_3 \\ \hline z_4 & z_5 & z_6 \\ \hline z_7 & z_8 & z_9 \\ \hline \end{array}$ | $\begin{array}{ c c c } \hline -1 & -1 & -1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$ | $\begin{array}{ c c c } \hline -1 & 0 & 1 \\ \hline -1 & 0 & 1 \\ \hline -1 & 0 & 1 \\ \hline \end{array}$ |
|  |   | <b>Prewitt</b>   |  |
|  | $\begin{array}{ c c } \hline -1 & 0 \\ \hline 0 & 1 \\ \hline \end{array}$  | $\begin{array}{ c c c } \hline -1 & -2 & -1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$ | $\begin{array}{ c c c } \hline -1 & 0 & 1 \\ \hline -2 & 0 & 2 \\ \hline -1 & 0 & 1 \\ \hline \end{array}$ |
|  | <b>Roberts</b>  | <b>Sobel</b>   |  |

图 10.14 一幅图像的  $3 \times 3$  区域 ( $z$  项是灰度值) 和用于计算标记点  $z_5$  处的梯度的不同模板

### 8.2.1 Marr-Hildreth边缘检测器

由于拉普拉斯算子的应用通常会放大图像的噪声，因此通常先平滑，再应用拉普拉斯算子。假设  $f(x, y)$  为

图像,  $h(x, y)$ 为高斯平滑函数

$$h(x, y) = -e^{-\frac{x^2+y^2}{2\sigma^2}}$$

则只需做一步操作完成平滑及边缘检测

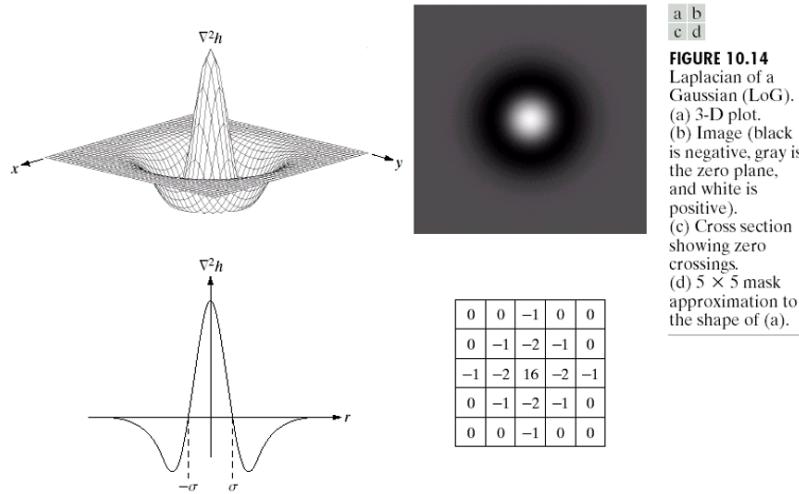
$$\nabla^2[f(x, y) * h(x, y)] = f(x, y) * [\nabla^2 h(x, y)]$$

其中

$$\nabla^2 h(x, y) = \frac{2}{\sigma^2} \left[ 1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2+y^2}{2\sigma^2}}$$

即高斯拉普拉斯(LoG)变换。

墨西哥草帽函数。(高斯函数的微分就是一种小波, 做微分后负号与负号抵消; 而且高斯函数有无穷阶导数)



得到Marr-Hildreth的边缘检测算法

1. 计算高斯拉普拉斯变换(LoG)
2. 找到零交叉点 (边缘正负变化的地方)

用高斯差分DoG滤波可以近似代替LoG,

$$DoG(x, y) = \frac{1}{2\pi\sigma_1^2} e^{-\frac{x^2+y^2}{2\sigma_1^2}} - \frac{1}{2\pi\sigma_2^2} e^{-\frac{x^2+y^2}{2\sigma_2^2}}, \sigma_1 > \sigma_2$$

以 $\sigma_1 : \sigma_2 = 1.6 : 1$ 来定, 则近似的 $\sigma$ 为

$$\sigma^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 - \sigma_2^2} \ln \left[ \frac{\sigma_1^2}{\sigma_2^2} \right]$$

### 8.2.2 Canny边缘检测器

目标:

- 低错误率: 边缘一个不落, 一个不多

- 边缘点应该被很好定位：标记的边缘点与真实边缘中心之间的距离最小
- 单一边缘点响应：对于真实的边缘点，检测器仅返回一个点，即真实边缘周围的局部最大数应该最小

算法步骤：

1. 用一个高斯滤波器平滑输入图像
2. 计算梯度幅值图像和角度图像（Canny为二阶梯度算子）

$$M(x, y) = \sqrt{g_x^2 + g_y^2}, \quad \alpha(x, y) = \arctan \left[ \frac{g_y}{g_x} \right]$$

3. 对梯度幅值图像应用非最大抑制

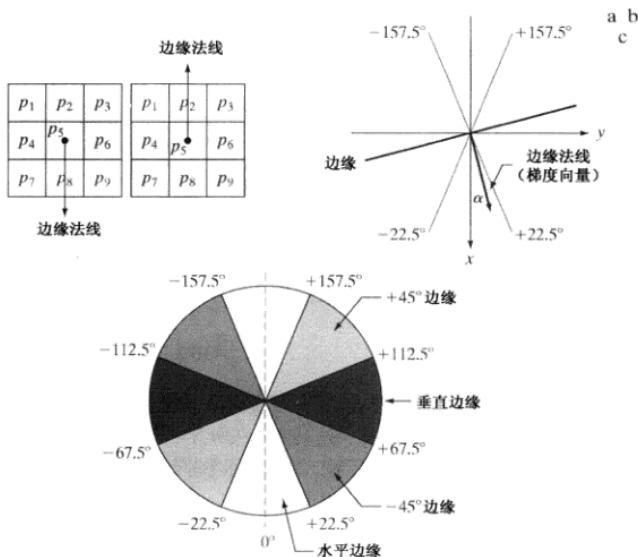


图 10.24 (a) 在一个  $3 \times 3$  邻域中，一个水平边缘(灰色)的两个可能方向；(b)一个水平边缘的边缘法线的方向角  $\alpha$  的值的范围(灰色)；(c)在一个  $3 \times 3$  邻域中，4 种类型的边缘方向的边缘法线的角度范围。每个边缘方向都有两个范围，以对应的灰色显示

(a) 寻找最接近  $\alpha(x, y)$  的方向  $d$

(b)  $g_N$  代表细化后的边缘

$$g_N(x, y) = \begin{cases} 0 & M(x, y) \text{ 的值至少小于沿 } d \text{ 的两个邻居之一} \\ M(x, y) & \text{否则} \end{cases}$$

4. 用双阈值处理和连接分析来检测并连接边缘

双阈值法：高阈值  $T_H$  和低阈值  $T_L$ ，比率为  $2 : 1$  或  $3 : 1$

$$g_{NN}(x, y) = g_N(x, y) \geq T_H \quad \text{强边缘}$$

$$g_{NL}(x, y) = g_N(x, y) \geq T_L \quad \text{弱边缘 (可能是边缘也可能不是)} + \text{强边缘}$$

$$g_{NH}(x, y) = g_{NL}(x, y) - g_{NN}(x, y) \quad \text{弱边缘}$$

用弱边缘补齐强边缘来获得完整边缘

- (a) 在 $g_{NN}(x, y)$ 中定位下一个未被访问的边缘像素 $p$
- (b) 在 $g_{NL}(x, y)$ 中用8连通方法连接到 $p$
- (c) 如果 $g_{NN}(x, y)$ 中所有非零标记都已经访问过，则跳到(d)，否则(a)
- (d) 将 $g_{NL}(x, y)$ 中未被标记为有效边缘的像素的所有像素置零。将 $g_{NL}(x, y)$ 中非零像素附加到 $g_{NN}(x, y)$

### 8.3 边缘连接

边界是封闭的边缘。

$$\text{边界检测} = \text{边缘检测} + \text{边缘连接}$$

- 边缘连接：两个端点只有在边缘强度和走向相近的情况下才能连接。如果像素 $(s, t)$ 在像素 $(x, y)$ 的邻域内且满足：

$$|M(x, y) - M(s, t)| \leq E, \quad |\alpha(x, y) - \alpha(s, t)| \leq A$$

则可以将 $(s, t)$ 与 $(x, y)$ 连接起来。

- 边界跟踪：依照角度搜索，用 $3 \times 3$ 区域平均值代替单像素点（避免噪声影响），称为虫

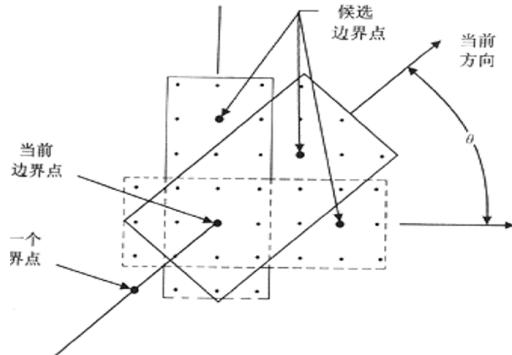


图 18.10 边界跟踪虫

- 区域处理：用多边形拟合算法

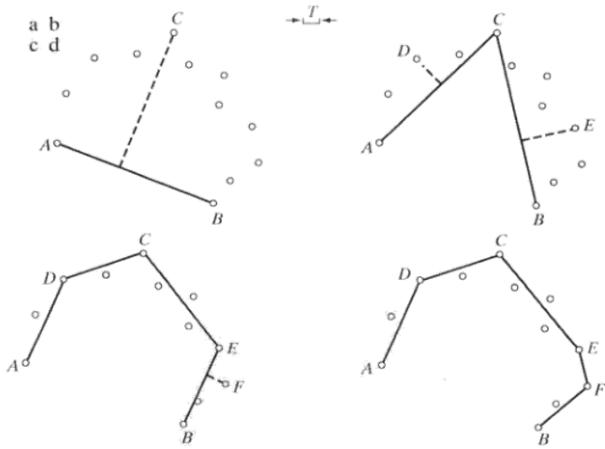


图 10.28 多边形拟合算法迭代说明

## 8.4 边界检测

用Hough变换进行边界检测：通过边界点找已知形状的目标。

### 8.4.1 直线检测问题

已知一组边缘点，找一条直线，使它通过最多边缘点。

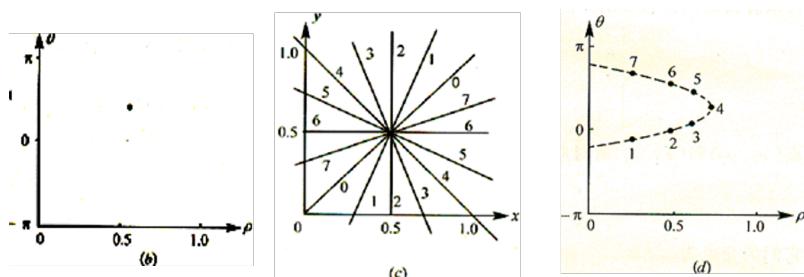
直线方程用极坐标表示

$$\rho = x \cos \theta + y \sin \theta$$

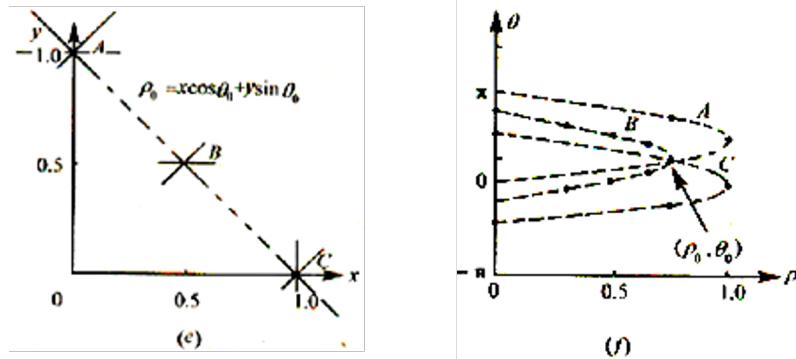
通过辅助角变换可得

$$\rho = A_0 \sin(\theta + \phi_0)$$

故可映射到 $\rho\theta$ 空间，其中每一个点是 $xOy$ 平面上的通过同一个点的一条线。



如果有一组位于由参数 $\rho_0$ 和 $\theta_0$ 决定的直线上的边缘点，则每个边缘点对应了 $\rho, \theta$ 空间的一条正弦曲线。所有这些曲线必然会交于点 $(\rho_0, \theta_0)$ ，因为这是它们共享的一条直线的参数。



故对于边缘点的直线拟合问题，即找一个使边缘点确定的正弦曲线相交最多的点 $(\rho_0, \theta_0)$ 。

可以建立 $\rho, \theta$ 空间的二维直方图来确定对于边缘点的最佳拟合直线参数 $(\rho_0, \theta_0)$ 。具体算法如下：

1. 对于每个边缘点 $(x, y)$ 。建立直线方程：

$$\rho = x \cos(\theta) + y \sin(\theta)$$

2. 假定 $\rho, \theta$ 的变化范围为 $\rho \in [\rho_{\min}, \rho_{\max}], \theta \in [\theta_{\min}, \theta_{\max}]$ ，建立 $A(\rho, \theta)$ 累加器
3. 给定 $\theta$ ，由原始方程确定 $\rho$ ，则 $A(\rho, \theta) += 1$
4. 对于所有的边缘点，执行上述步骤，找出最大的 $A(\rho, \theta)$ ，即为所求

#### 8.4.2 其他检测问题

换成圆/椭圆也一样

$$(x - x_0)^2 + (y - y_0)^2 = r^2$$

在参数空间建立3D累加数组 $A$ ，让 $(x_0, y_0)$ 变化，计算 $r$ 。

## 8.5 阈值处理

阈值/门限处理模型

$$T = T[x, y, p(x, y), f(x, y)]$$

如果

$$g(x, y) = \begin{cases} 1 & f(x, y) > T \\ 0 & f(x, y) \leq T \end{cases}$$

当 $T$ 为适用于整个图像的常数时，则该公式给出的处理为全局阈值处理， $f(x, y) > T$ 的任何点为一个对象点，否则该点为背景点。如果取决于邻域 $p(x, y)$ ，则为局部阈值处理。如果取决于 $(x, y)$ 本身，则为动态阈值处理/自适应阈值处理。

- 噪声太大时会将直方图中的多个峰融合：先做平滑

- 背景光照不均匀时阈值处理方法：

1. 直接矫正方法：用恒定灰度的平坦表面成像获得光照模式，用相反的模式与图像相乘来矫正

2. 采用顶帽变换来获得全局阴影模式
3. 使用可变阈值近似处理非均匀性

### 8.5.1 基本全局阈值处理

试探法：

1. 为全局阈值  $T$  选择一个初始估计值
2. 用  $T$  分割图像，生成两组像素：  $G_1$  由所有灰度值大于  $T$  的像素组成，而  $G_2$  由所有灰度值小于或等于  $T$  的像素组成。
3. 对区域  $G_1$  和  $G_2$  中的所有像素计算平均灰度值  $m_1$  和  $m_2$
4. 计算新的门限值：  $T = 1/2(m_1 + m_2)$
5. 重复步骤2到4，直到逐次迭代所得的  $T$  值之差小于实现定义的参数  $\Delta T$

### 8.5.2 Otsu最佳全局阈值处理

将其看为一个分类问题，用贝叶斯决策，这是最佳的方法。

1. 归一化直方图  $p_i = n_i/(NM)$
2. 给定阈值  $T(k)$ ，将其分为左右两个部分  $C_1$  和  $C_2$
3. 求出两块的条件概率密度  $p_i/P_1(k)$  和  $p_i/P_2(k)$
4. 定义归一化度量指标  $\eta = \frac{\sigma_B^2}{\sigma_G^2}$   
其中  $\sigma_G^2$  为全局方差

$$\sigma_G^2 = \sum_{i=0}^{L-1} (i - m_G)^2 p_i$$

$\sigma_B^2$  为类间方差

$$\begin{aligned} \sigma_B^2 &= P_1(m_1 - m_G)^2 + P_2(m_2 - m_G)^2 \\ &= P_1 P_2 (m_1 - m_2)^2 \\ &= \frac{(m_G P_1 - m)^2}{P_1(1 - P_1)} \end{aligned}$$

5. 基本结论是  $m_1$  和  $m_2$  相隔越大，则  $\sigma_B^2$  越大；  $\eta$  是分割的可分性度量， $\sigma_B^2$  越大，则  $\eta$  越大。进而有最佳阈值

$$\sigma_B^2(k^*) = \max_{0 \leq k \leq L-1} \sigma_B^2(k)$$

### 8.5.3 预处理

用图像平滑来改善全局阈值处理：但难以处理单峰

用边缘来改进全局阈值处理：梯度算法能够区分边缘区域与平坦区域；拉普拉斯算子可以确定给

定的像素在边缘的亮一边还是暗一边。局部阈值法：

$$s(x, y) = \begin{cases} 0 & \nabla f < T \\ + & \nabla f \geq T, \nabla^2 f \geq 0 \\ - & \nabla f \geq T, \nabla^2 f < 0 \end{cases}$$

1. 计算  $f(x, y)$  的梯度图或者拉普拉斯绝对值图
2. 给定一个阈值
3. 用步骤2)的结果做阈值处理（通常取第  $n$  个百分比），产生二值图像  $g_T(x, y)$ （可以是梯度二值图和拉普拉斯绝对值图二值图取“或”的组合）
4. 计算  $h(x, y) = f(x, y) \cdot g_T(x, y)$ ，计算  $h(x, y)$  的直方图
5. 在4)直方图的基础上，用Otsu方法做图像分割

#### 8.5.4 多阈值处理

在  $K$  个类的情况下，同样可以定义类间方差，来计算最大的归一化度量指标

#### 8.5.5 可变阈值处理

通过图像分块的阈值处理，因为块内的光照近似均匀

### 8.6 基于区域的分割

$f(x, y)$  为图像， $S(x, y)$  为种子阵列，种子处为1，其他为0。 $Q$  为位置  $(x, y)$  的属性，基于8连通的区域生长算法为：

1. 在  $S(x, y)$  中找连通分量，并将连通分量腐蚀为一个像素；将找到的所有这种像素标记为1， $S$  的其他像素标记为0
2. 计算

$$f_Q(x, y) = \begin{cases} 1 & (x, y) \text{ 处属性 } Q \text{ 为真} \\ 0 & \text{其他} \end{cases}$$

3. 分割后图像  $g$  为：把  $f_Q$  中与种子点8连通的所有1值点添加到  $S$  中的每个种子点
4. 使用不同的区域标记出  $g$  中的每个连通分量
  - 区域生长法
  - 区域分离合并法

### 8.7 基于形态学分水岭的分割

基本思想：

1. 将梯度值图像看成一幅地形图，梯度值对应海拔高度，图像中不同梯度值的区域就对应于山峰和山谷间盆地。
2. 设想在各个局部极小值点的位置打一个洞，让水以均匀上升速率从洞中涌出，从低到高淹没整个地形。
3. 水位逐渐升高漫过盆地，当相邻两个盆地的水即将合并时，这时在两个盆地间建坝拦截。
4. 此过程将图像划分为许多个山谷盆地，分水岭就是分隔这些盆地的堤坝。

水坝构造：以二值图像为基础，使用形态膨胀的方法分离二元点集构造水坝。

- $M_1, M_2$  表示在两个区域极小值中包含的点的坐标集合
- 溢出的第  $n - 1$  阶段与  $M_1, M_2$  联系的处于汇水盆地中的点集
- $C[n - 1]$  表示  $C_{n-1}(M_1)$  和  $C_{n-1}(M_2)$  并集
- $q$  为第  $n$  步聚合后的连通分量
- $q \cap C[n - 1]$  可从中提取第  $n - 1$  步的两个连通分量

变量声明：

- $M_1, M_2, \dots, M_R$  表示局部最小值点的坐标的集合
- $C(M_i)$  表示与局部最小值  $M_i$  相联系的汇水盆地内的点的集合
- $T[n] = \{(s, t) \mid g(s, t) < n\}$  表示位于平面  $g(x, y) = n$  下方的点的集合
- $C_n(M_i)_R = C(M_i) \cap T[n]$  表示第  $n$  阶段汇水盆地  $i$  中被淹没的点的集合
- $C[n]$  表示第  $n$  阶段汇水盆地被水淹没的点的集合
- $C[max + 1] = \bigcup_{i=1}^R C(M_i)$  表示所有汇水盆地的集合
- $C[n - 1] \subset C[n] \subset T[n]$  即  $C[n - 1]$  中每个连通分量都恰好是  $T[n]$  的一个连通分量

分水岭分割算法：

- 初始化：  $C[min + 1] = T[min + 1]$
- 递归： 根据  $C[n - 1]$  求得  $C[n]$  的过程如下：
  - (1) 遇到新的最小值，符合条件(a)，将  $q$  并入  $C[n - 1]$  构成  $C[n]$
  - (2)  $q$  位于某些局部最小值构成的汇水盆地中，符合条件(b)
  - (3) 遇到全部或部分分离汇水盆地的山脊线，符合条件(c)，在  $q$  内构造水坝，即得到分水线

其中  $Q$  代表  $T[n]$  中连通分量的集合，对每个连通分量  $q \in Q[n]$ ，有 3 种可能：

- (a)  $q \cap C[n - 1]$  为空
- (b)  $q \cap C[n - 1]$  包含  $C[n - 1]$  中的一个连通分量
- (c)  $q \cap C[n - 1]$  包含  $C[n - 1]$  多于一个的连通分量

分水岭分割算法的缺点：

- 对图像中的噪声极为敏感。由于输入图像往往是图像梯度，原始图像中的噪声能直接恶化图像的梯度，造成分割的轮廓偏移。

- 易于产生过度分割。由于受噪声和平坦区域内部细密纹理的影响，导致局部极值过多，在后续分割中出现大量的细小区域。

解决方案：

- 分割前预处理
- 分割时添加约束
- 分割后对图像进行再处理

## 8.8 分割中运动的作用

- 运动是人类和动物使用的用于将重要对象从不相关的背景细节中提取出的强有力的线索。
- 用于成像应用领域，如机器人应用，自主导航和动态视觉分析。

基本方法：

- 差值图像
- 分别检测两帧图像 $f(x, y, t_1)$ 和 $f(x, y, t_2)$ 在时间 $t_1$ 和 $t_2$ 时的变化的最简单的方法是将两幅图像逐个像素的进行对比。
- 这个过程将得到一幅差值图像

$$d_{ij}(x, y) = \begin{cases} 1 & |f(x, y, t_i) - f(x, y, t_j)| > T \\ 0 & \text{其他} \end{cases}$$

其中 $T$ 是一个特定的门限。从定义可以看到：如果两幅图像中对应的坐标上的灰度差有相当的不同，则 $d_{ij}(x, y)$ 具有值1，差异程度取决于事先确定的门限 $T$

- 假设所有的图像具有相同的尺寸大小，则差值图像也具有相同的尺寸
- 在动态处理过程中 $d_{ij}(x, y)$ 所有值为1的像素被认为是对象运动的结果
- 但是实际上，这些1值常常是由于噪声造成的，对于这种情况，我们可以使用一些简单的方法来除去他们

考虑三种差异累计图(ADI)：绝对ADI，正ADI，负ADI

$$A_k(x, y) = \begin{cases} A_{k-1}(x, y) + 1 & |f(x, y, 1) - f(x, y, k)| > T \\ A_{k-1}(x, y) & \text{其他} \end{cases}$$

$$P_k(x, y) = \begin{cases} P_{k-1}(x, y) + 1 & f(x, y, 1) - f(x, y, k) > T \\ P_{k-1}(x, y) & \text{其他} \end{cases}$$

$$N_k(x, y) = \begin{cases} N_{k-1}(x, y) + 1 & f(x, y, 1) - f(x, y, k) < -T \\ N_{k-1}(x, y) & \text{其他} \end{cases}$$

# 9 表示与描述

图像表示与描述有两种方法:

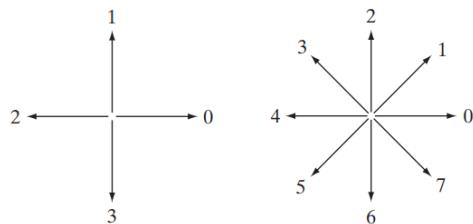
- 外部描述: 形状
- 内部描述: 颜色与纹理

## 9.1 表示方法

### 9.1.1 链码

链码用于表示由顺次连接的具有指定长度和方向的直线段组成的边界线

链码种类: 4向链码和8向链码



常见的问题:

1. 得到的链码太长;
2. 噪声或边界缺陷的影响;

解决方案: 选择更大间隔的网格对边界进行重新采样。

- 边界链码依赖于起始点 (通过归一化可一致, 即使链码最小)
- 用差分码 (逆时针旋转的方向数) 可以将链码旋转归一化

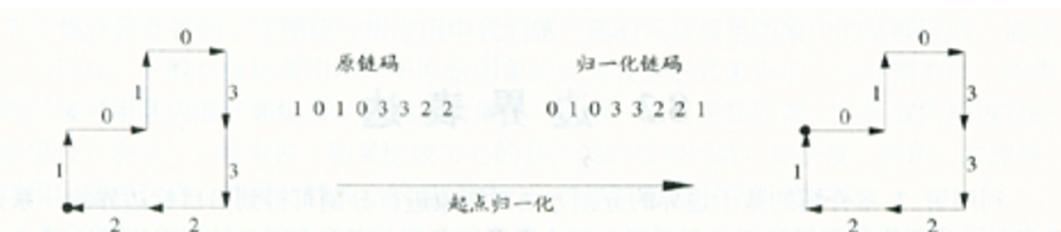


图 8.2.2 链码的起点归一化

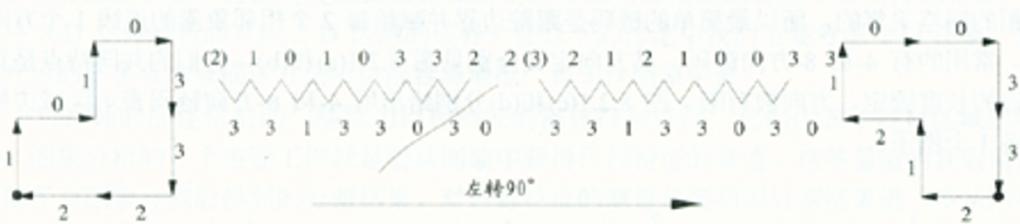


图 8.2.3 链码的旋转归一化 (利用一阶差分)

### 9.1.2 多边形近似表达方法

- 基于收缩的最小周长多边形法：将原边界看作有弹性的线，在像素的约束下将线拉紧
- 基于聚合的最小均方误差线段逼近法：每个顶点到边的垂线距离小于 $\delta$
- 基于分裂的最小均方误差线段逼近法

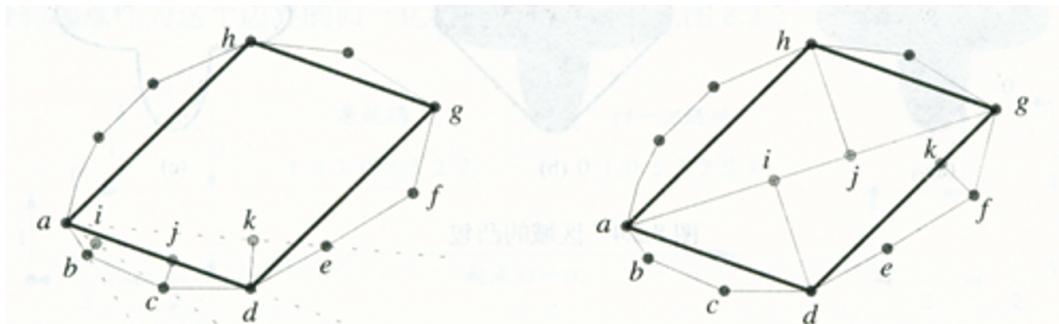
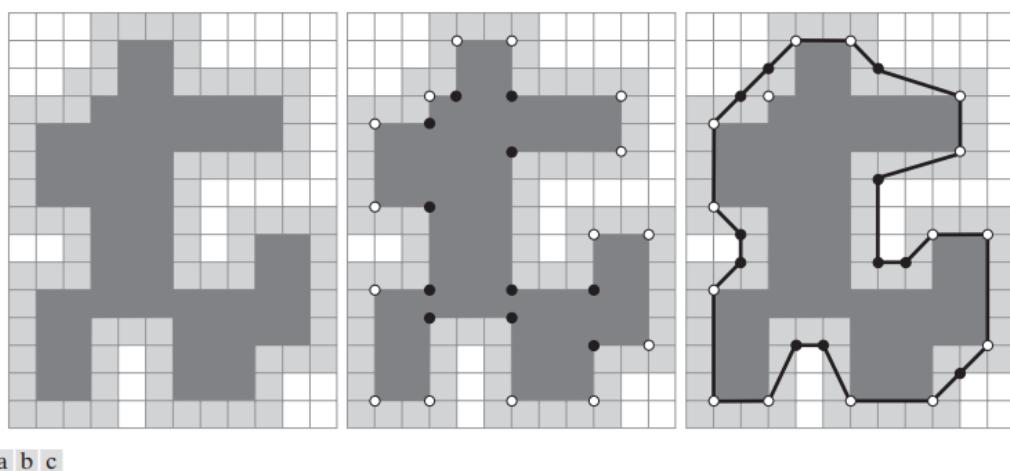


图 8.2.6 聚合逼近多边形

图 8.2.7 分裂逼近多边形

基于收缩的最小周长多边形法（MPP算法）：围成一条数字边界的单元集合称为单元组合体。令W、B分别表示凸顶点和镜像凹顶点的集合。

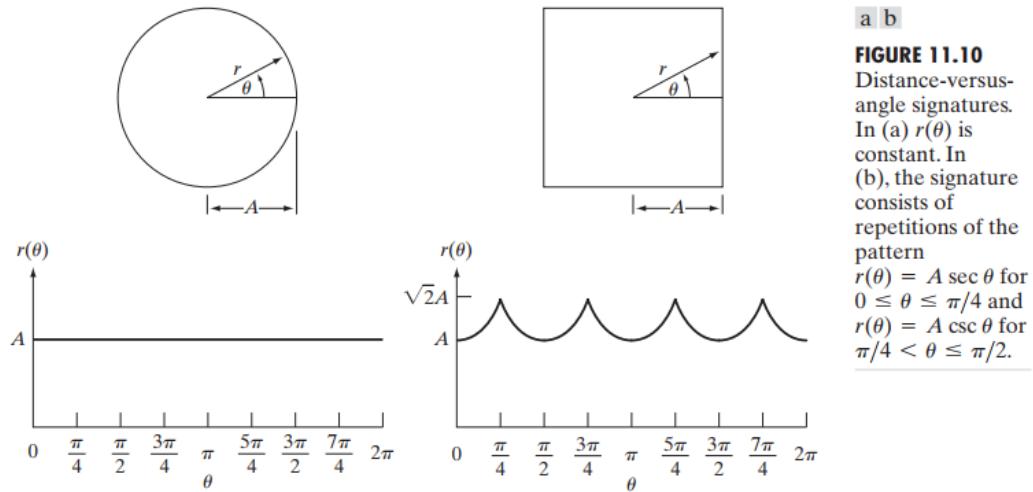
- 由简单连接的单元组合体为边界的MPP是非自相交的。
- MPP的每个凸顶点都是一个W顶点，但并非边界的每个W顶点都是MPP的一个顶点。
- MPP的每个镜像凹顶点都是一个B顶点，但并非边界的每个B顶点都是MPP的一个顶点。
- 所有B顶点，要么在MPP上，要么在MPP外；所有W顶点，要么在MPP上，要么在MPP内
- 单元组合体包含的顶点序列的最左上角的顶点，总是MPP的一个W顶点。



**FIGURE 11.7** (a) Region (dark gray) resulting from enclosing the original boundary by cells (see Fig. 11.6). (b) Convex (white dots) and concave (black dots) vertices obtained by following the boundary of the dark gray region in the counterclockwise direction. (c) Concave vertices (black dots) displaced to their diagonal mirror locations in the outer wall of the bounding region; the convex vertices are not changed. The MPP (black boundary) is superimposed for reference.

### 9.1.3 标记图方法

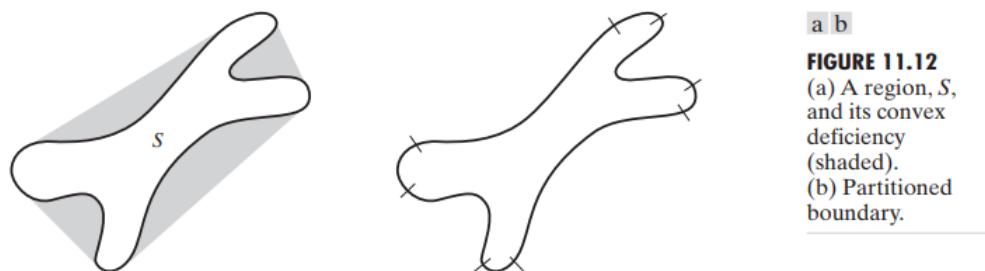
标记图是一种一维函数的边界表达方法



- 不受目标平移影响，但受目标尺度变换和旋转的影响
- 尺度变换影响可以通过将最大幅值归一化来解决
- 旋转的影响解决方法
  - 选择离重心最远的点为标记起点
  - 求出边界的主轴，以主轴上离重心最远的点作为标记起点

### 9.1.4 边界线段

将边界分段可以减小边界的复杂度。



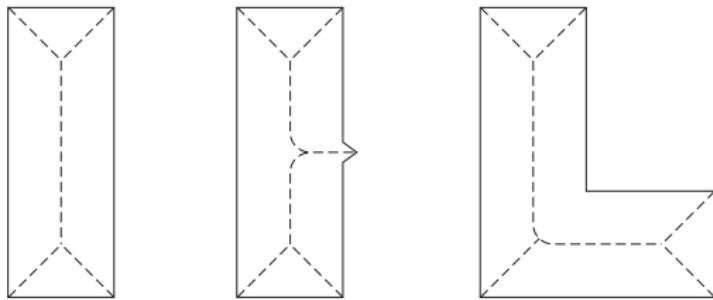
### 9.1.5 骨架

表达平面区域结构形状的一种方法；此方法可以用细化算法实现。

a b c

FIGURE 11.13

Medial axes  
(dashed) of three  
simple regions.



集合 $A$ 的骨架 $S(A)$ 定义为

$$S(A) = \{z \mid A \text{内最大圆盘}(D)_z \text{的圆心}\}$$

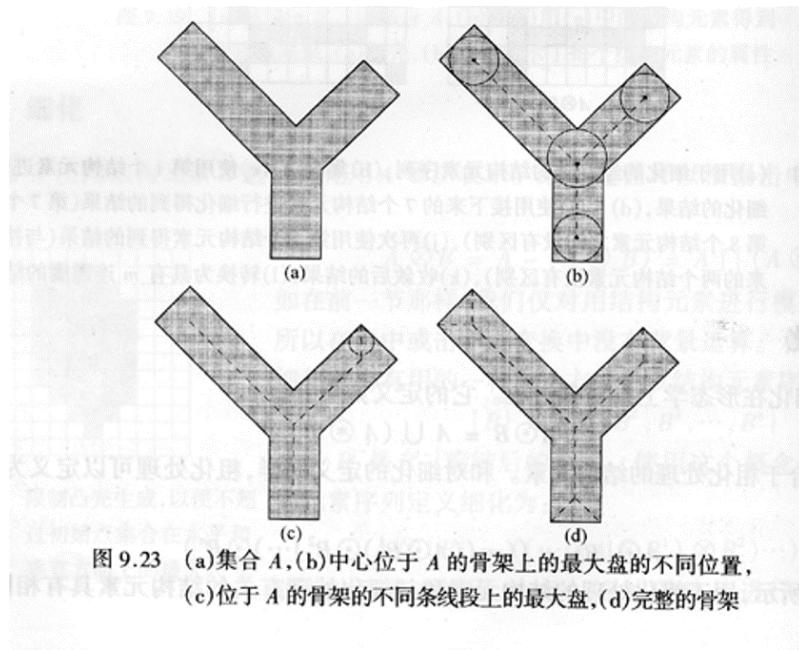


图 9.23 (a)集合  $A$ , (b)中心位于  $A$  的骨架上的最大盘的不同位置,  
(c)位于  $A$  的骨架的不同条线段上的最大盘,(d)完整的骨架

## 9.2 边界描绘子

### 9.2.1 简单的描述符

1. 边界的长度;
2. 边界的直径: 边界上相隔最远两点的距离。( $D$ 可以是任意距离度量)

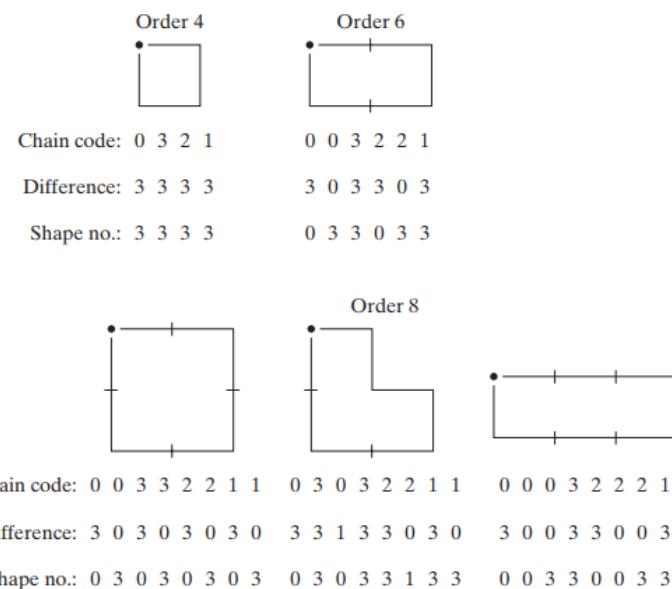
$$Diam(B) = \max_{i,j} D(p_i, p_j), p_i \in B, p_j \in B$$

3. 曲率: 斜率的变化率。通常采用相邻边界线段的斜率差。当沿顺时针方向沿着边界运动, 当顶点 $p$ 斜率变化量为非负的时候, 称这一点属于凹线段; 否则, 称 $p$ 属于凸线段。

## 9.2.2 形状数

形状数是基于链码的一种边界形状描述符。一个边界形状描述是其链码的差分码中值最小的一个序列。形状数长度称为阶数。

**FIGURE 11.17**  
All shapes of  
order 4, 6, and 8.  
The directions are  
from Fig. 11.3(a),  
and the dot  
indicates the  
starting point.



## 9.2.3 傅里叶描绘子

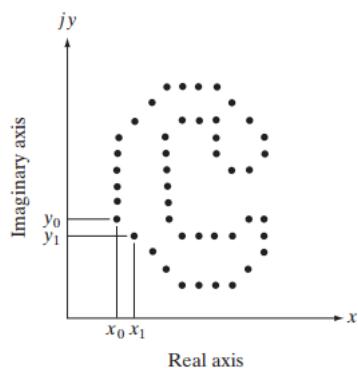
将2D问题转化为1D问题，将边界点的坐标对表示成一个复数

$$s(k) = x(k) + jy(k), k = 0, 1, \dots, K-1$$

对  $s(k)$  的傅氏变换为

$$a(u) = \frac{1}{K} \sum_{k=0}^{K-1} s(k) e^{-j2\pi uk/K}, u = 0, 1, \dots, K-1$$

复系数  $a(u)$  称为边界的描绘子。



傅氏反变换确定边界的重建:

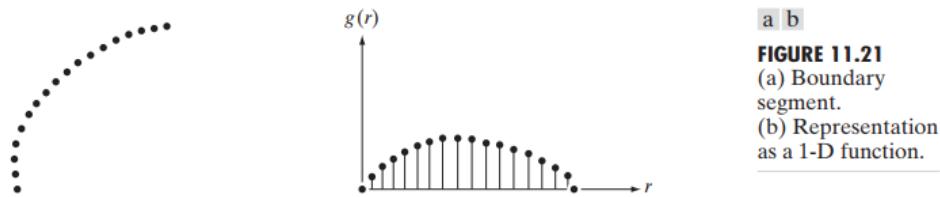
$$s(k) = \sum_{u=0}^{K-1} a(u) e^{j2\pi uk/K}, k = 0, 1, \dots, K-1$$

通过有限重建构造近似边界

$$\hat{s}(k) = \sum_{u=0}^{P-1} a(u) e^{j2\pi uk/K}, k = 0, 1, \dots, K-1, K > P$$

#### 9.2.4 统计矩

统计矩(statistical moment)用于刻画边界线段的特征波形。



1. 将上述曲线看作一维函数  $g(r)$
2. 将  $g$  的振幅看作离散随机变量  $v$ , 并形成直方图  $p(v_i)$ , 其中  $i = 0, 1, \dots, A-1$
3. 定义  $n$  阶中心矩

$$\mu_n(v) = \sum_{i=0}^{A-1} (v_i - m)^n p(v_i)$$

其中  $m = \sum_{i=0}^{A-1} v_i p(v_i)$ ,  $m$  是  $v$  的均值

另一种统计矩:

1. 将  $g(r)$  归一化为单位面积的函数, 并作成直方图, 即将  $g(r_i)$  作为产生值  $r_i$  的概率
2. 定义  $n$  阶中心矩为

$$\mu_n(r) = \sum_{i=0}^{K-1} (r_i - m)^n g(r_i)$$

其中  $m = \sum_{i=0}^{K-1} r_i g(r_i)$ ,  $m$  是  $r$  的均值

### 9.3 区域描绘子

#### 9.3.1 简单的描述符

1. 区域的面积: 图象中对象区域的面积可以看作区域像素的总和。通常关心一些不变量, 如区域致密性:

$$(周长)^2 / \text{面积}$$

## 2. 区域的重心

$$\begin{cases} \bar{x} = \frac{1}{A} \sum_{(x,y) \in \mathcal{R}} x f(x, y) \\ \bar{y} = \frac{1}{A} \sum_{(x,y) \in \mathcal{R}} y f(x, y) \end{cases}$$

其中， $\mathcal{R}$ 代表一个区域， $A$ 代表区域面积。

## 3. 区域灰度/密度：常用的区域灰度特征有目标灰度（或颜色分量）的最大值、最小值、中值、平均值、方差以及高阶矩等统计量。

### 9.3.2 拓扑描述符

欧拉数

$$E = C - H$$

其中 $C$ 为区域内连通组元数， $H$ 为区域内孔数。

将区域的网络进行目标区域的分类，可以分为顶点数 $V$ ，边数 $Q$ ，面数 $F$ ，其欧拉公式为：

$$V - Q + F = C - H = E$$

### 9.3.3 纹理

- 纹理就是由纹理基元按某种确定的规律或某种统计规律排列而成
- 纹理分为确定性纹理和随机性纹理

区域的纹理主要度量区域的平滑度、粗糙度和规律性。描述纹理的方法主要有三种：统计方法、结构方法和频谱方法。下文着重考虑统计方法。

区域灰度直方图的统计矩（ $n$ 阶中心矩）：

$$\mu_n(z) = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i)$$

其中 $p(z_i)$ 为归一化直方图

光滑度描述子

$$R = 1 - \frac{1}{1 + \sigma^2(z)}$$

若 $R = 0$ 则表示平滑（区域平坦）， $R = 1$ 代表不平滑

其他度量

- 一致性

$$U = \sum_{i=0}^{L-1} p^2(z_i)$$

- 熵

$$p = - \sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$$

共生矩阵：设 $S$ 为目标区域 $R$ 中具有特定空间联系的像素对的集合，则共生矩阵 $G$ 可定义为：

$$G(g_1, g_2) = \#\{(x_1, y_1), (x_2, y_2)] \in S \mid f(x_1, y_1) = g_1 \& f(x_2, y_2) = g_2\}$$

基于共生矩阵的描绘子

$$\begin{aligned} m_x &= \sum_i^K i \sum_j^K p_{ij}, & m_y &= \sum_j^K j \sum_i^K p_{ij} \\ \sigma_x^2 &= \sum_i^K (i - m_x)^2 \sum_j^K p_{ij}, & \sigma_y^2 &= \sum_j^K (j - m_y)^2 \sum_i^K p_{ij} \\ P(i) &= \sum_i^K p_{ij}, & P(j) &= \sum_j^K p_{ij} \\ m_x &= \sum_i^K i P(i), & m_y &= \sum_i^R j P(j) \\ \sigma_x^2 &= \sum_i^K (i - m_x)^2 P(i), & \sigma_y^2 &= \sum_j^R (j - m_y)^2 P(j) \end{aligned}$$

- 最大概率  $\max_{i,j}(p_{ij})$
- 对比度（元素差异的 $k$ 阶矩）

$$\sum_i^K \sum_j^K (i - j)^k p_{ij}$$

- 同质性： $G$ 对角分布的紧密性

$$\sum_i^K \sum_j^K \frac{p_{ij}}{1 + |i - j|}$$

- 一致性：

$$\sum_i^K \sum_j^K p_{ij}^2$$

- 熵：

$$-\sum_i^K \sum_j^K p_{ij} \log_2 p_{ij}$$

- 相关性：

$$\sum_i^K \sum_j^K \frac{(i - m_x)(j - m_y)p_{ij}}{\sigma_x \sigma_y}$$

二阶函数的矩：对于二维连续函数 $f(x, y)$ ， $(p+q)$ -阶矩定义为

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy$$

## 9.4 主分量描绘

主分量分析(PCA)、Hotelling变化、特征变量变换、K-L变换

协方差矩阵

$$C_x = \mathbb{E}((x - \bar{x})(x - \bar{x})^T) \approx \frac{1}{L} \sum_{l=1}^L x_l x_l^T - \bar{x} \bar{x}^T$$

其中  $C_x$  是半正定的实对称矩阵。

将图像堆积成列向量

$$\mathbf{x} = [f(0,0), f(0,1), \dots, f(0,N-1), f(1,0), f(1,1), \dots, f(1,N-1), f(N-1,0), f(N-1,1), \dots, f(N-1,N-1)]^T$$

对  $C_x$  进行特征分析，得到一组实非负特征值  $\lambda_1, \lambda_2, \dots, \lambda_r$  及对应的特征向量  $V_1, V_2, \dots, V_r$ 。

令  $T = [V_1 \ V_2 \ \dots \ V_r]$ ，则  $C_x T^T = [\lambda_i] T^T$ ，做变换  $y = T(x - \bar{x})$ ，则

$$\begin{aligned} C_y &= \mathbb{E}(yy^T) = \mathbb{E}(T(x - \bar{x})(x - \bar{x})^T T^T) \\ &= T \mathbb{E}((x - \bar{x})(x - \bar{x})^T) T^T \\ &= T C_x T^T = T[\lambda_i] T^T \\ &= [\lambda_i] = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_N \end{bmatrix} \end{aligned}$$

即  $y$  各元素互不相关， $T$  为去相关变换， $\lambda_k$  为变换后第  $k$  个分量  $y_k$  的方差。任何向量  $x$  都可以通过  $y$  重构

$$x = T^T y + \bar{x}$$

将特征根从大到小排列，记为  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$ ，同时将特征向量做相应排序，记为  $V_1, V_2, \dots, V_r$ 。

取  $1 < M < r$ ，使得  $\sum_{i=1}^M \lambda_i / \sum_{i=1}^r \lambda_i > \alpha$  (如  $\alpha = 95\%$ )

又令  $T_M = [V_1 \ V_2 \ \dots \ V_M]^T$ ， $\hat{y} = T_M(x - \bar{x})$ ，而  $x$  依然可以通过

$$\bar{x} = T_M^T \hat{y} + \bar{x}$$

来近似重构，其均方误差为

$$\begin{aligned}MSE &= \mathbb{E}((x - \hat{x})^T(x - \hat{x})) \\&= \mathbb{E}\left(\left(\sum_{i=M+1}^r y_i V_i\right)^T \left(\sum_{i=M+1}^r y_i V_i\right)\right) \\&= \sum_{i=M+1}^r \mathbb{E}(y_i^2) \\&= \sum_{i=M+1}^r \lambda_i\end{aligned}$$

基于矩阵的奇异值分解(SVD)， $\sqrt{\lambda_i}$ 为 $A$ 的奇异值，有推论 $u = Av\Lambda^{-1/2}$ ，即可取 $AA^T$ 与 $A^TA$ 中维数较小的矩阵计算特征向量。

可以将奇异值分解和PCA应用于多光谱图像的压缩。