

第一章 绪论

图像：二维函数 $f(x, y)$ ，其中 x 与 y 为平面坐标，幅值 f 称为该点处的强度或灰度。

数字图像领域的实例（按照能量级分段运用）：伽马射线成像、X射线成像、紫外波段成像、可见光及红外波段成像、微波波段成像、无线电波段成像、其它方式成像。

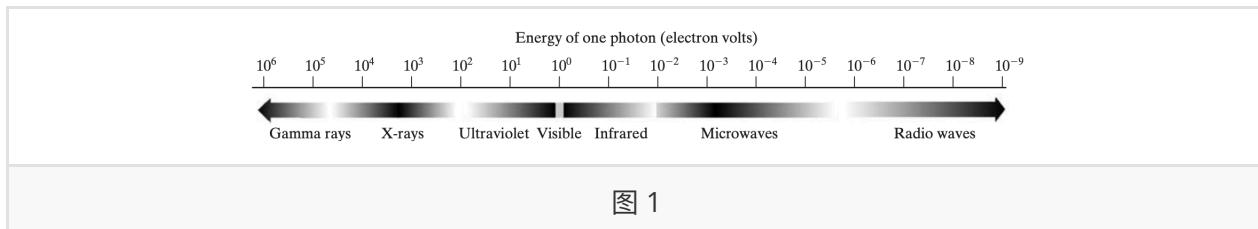


图 1

第二章 数字图像基础

2.1 视觉感知要素

眼睛由三层膜构成：角膜+虹膜外壳；脉络膜；视网膜。

视网膜上有光感受器，分为两类：

- 锥状体 (cones): 对颜色高度敏感，锥状体视觉称为白昼视觉（亮视觉）；
- 杆状体 (rods): 无彩色感觉，对低照明度敏感，杆状体视觉称为微光视觉（暗视觉）；
- 锥状体和杆状体的分布使眼睛存在“盲点”，如图 2 所示。
- 使用强度为 I 的可变光源从后方照向漫反射体，在视野中增加一个照射分量 ΔI ，形成一个持续时间很短的闪烁，当 ΔI 逐渐加强后，目标会给出肯定的答复指出一个可觉察的变化，则韦伯比 定义为 $\Delta I_c / I$ ，其中， ΔI_c 为背景为 I 时可辨别照明增量的 50%。

韦伯比大说明亮度辨别能力“较差”，反之说明亮度辨别能力“较好”。由图 3 可知，在低照明度情况下，视觉由杆状体执行；在高照明度情况下，视觉由锥状体执行。

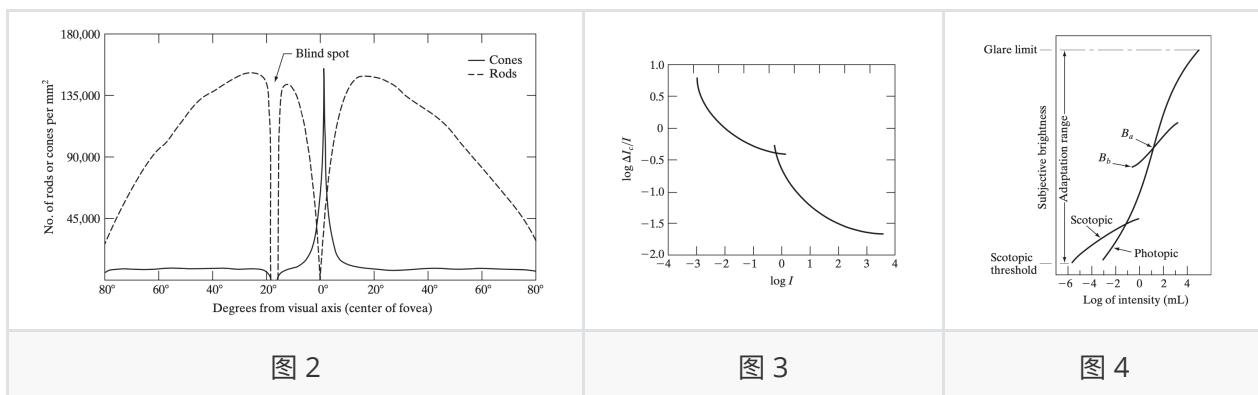


图 2

图 3

图 4

主观亮度 (subjective brightness): 由视觉系统感知的亮度，又称为感知亮度 (perceived brightness)；实验数据表明其是进入人眼光强的对数函数；马赫带 & 同时对比现象表明其不是光强的简单函数。

亮度适应现象：视觉系统不能同时在一个范围内工作，其通过改变整个灵敏度完成这一变动，视觉当前灵敏度级别称为亮度适应级别，如图 4 所示。

错觉 是人类视觉系统的另一特性，但暂未被完全了解。

2.2 光和电磁波谱

单色光（无色光）：没有颜色的光称为单色光（或无色光），其唯一属性是其强度（或大小）。

灰度级：表示单色光的强度，从黑到白的单色光度量值范围称为灰度级，单色图像通常称为灰度图像。

三个基本量表示彩色光源的质量

- **发光强度**：从光源流出的能量总量，用瓦特 (W) 度量；
- **光通量**：观察者从光源感受的能量，用流明数 (lm) 度量；
- **亮度**：是光感知的主观描绘子，实际上不能度量。

2.3 图像的感知和获取

将照射能量变换为数字图像的三种主要传感器配置（图 5）

- 单个成像传感器：传感器线性运动、胶片在圆柱体上沿轴转动；
- 条带传感器：传感器线性运动，或三维物体线性运动（进入环传感器）；
- 阵列传感器：静止即可（如数字摄像机）。

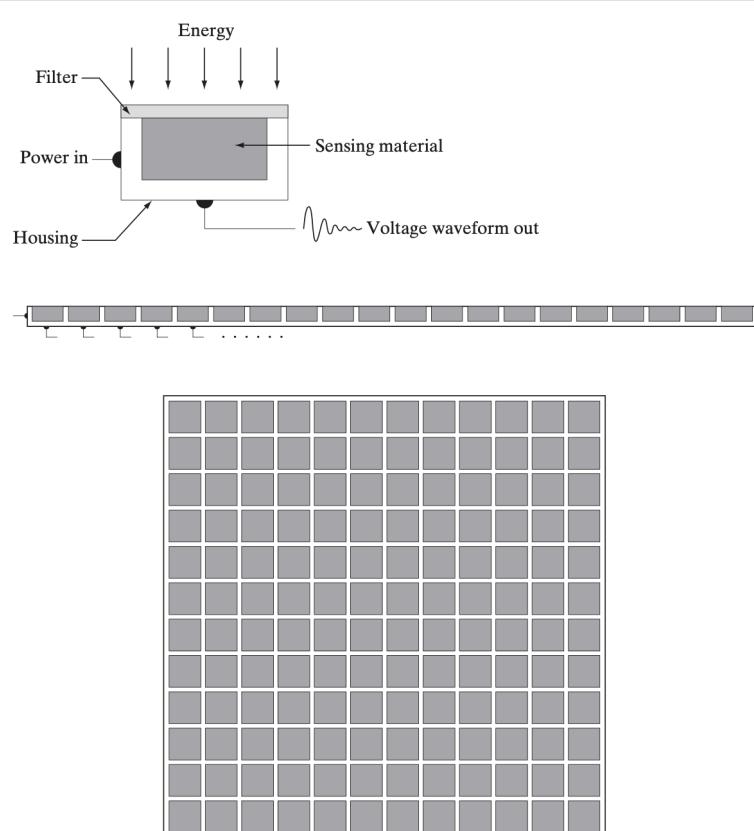


图 5

简单的图像形成模型：形如 $f(x, y)$ 的代表亮度值的二维函数表示图像（正标量），其中 $0 < f(x, y) < +\infty$ ； $f(x, y)$ 可以用两个分量来表征：入射到被观察场景的光源照射总量 $i(x, y)$ 与场景中物体反射的光照总量 $r(x, y)$ 。两个函数的乘积为 $f(x, y)$ ，即

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

其中, $0 < i(x, y) < +\infty$, $0 < r(x, y) < 1$; 其中 $r(x, y) = 0$ 为全吸收, $r(x, y) = 1$ 为全反射。 $i(x, y)$ 的性质取决于照射源; 而 $r(x, y)$ 的性质取决于成像物体的特性。

灰度级离散化: 令单色图像在任何坐标 (x_0, y_0) 处的强度 (灰度) 表示为

$$\mathcal{L} = f(x_0, y_0)$$

则 \mathcal{L} 满足如下式:

$$L_{min} \leq \mathcal{L} \leq L_{max}$$

理论上, 要求 L_{min} 为正而 L_{max} 为有限值。事实上 (为简化), 可令 $L_{min} = i_{min}r_{min}$, $L_{max} = i_{max}r_{max}$ 。那么, 区间 $[L_{min}, L_{max}]$ 称为灰度级 (或强度级)。实际上, 令该区间为 $[0, L - 1]$, 其中 $\mathcal{L} = 0$ 为黑色, $\mathcal{L} = L - 1$ 在灰度级中为白色; 所有中间值是从黑色到白色之间变化的灰度色调。

2.4 图像取样与量化

将连续的感知数据转化为数字形式需要两种处理: 取样与量化, 如图 6 所示。

- 对坐标值进行数字化称为取样;
- 对幅值进行数字化称为量化。

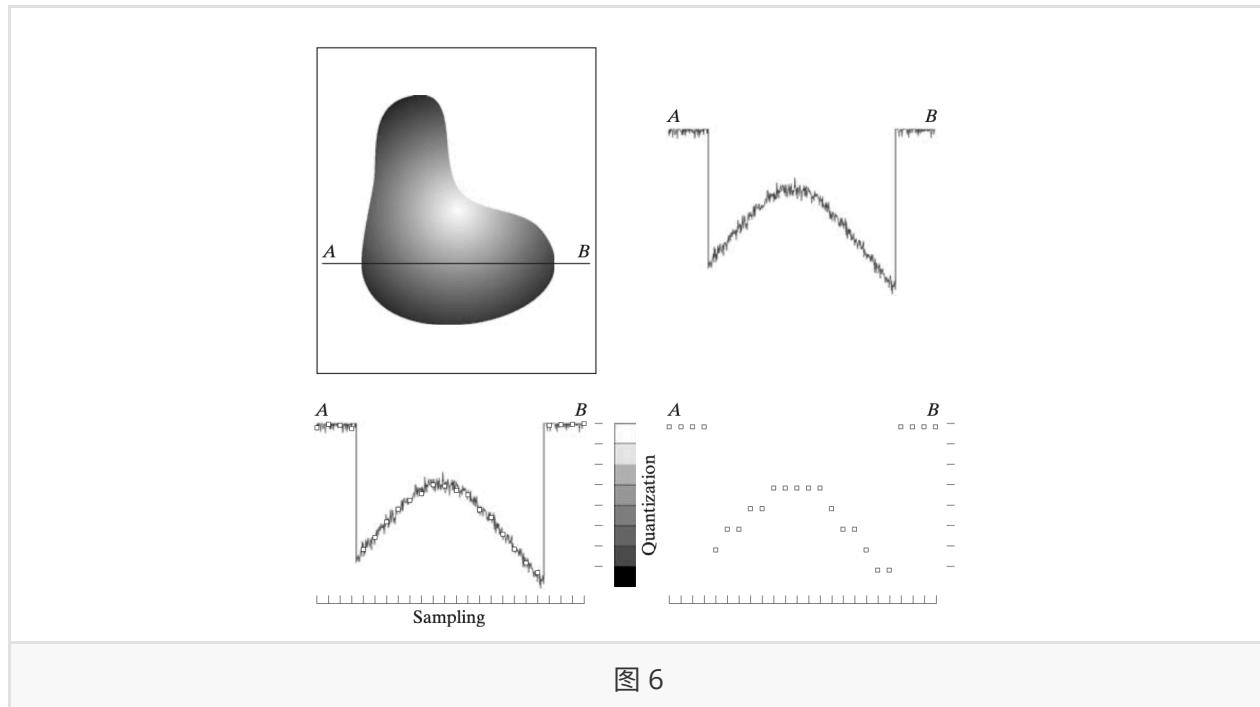


图 6

数字图像表示: 令 $f(s, t)$ 表示一幅具有两个连续变量 s 和 t 的连续图像函数。通过取样和量化, 我们可把该函数转化为数字图像。假设我们把这幅连续图像取样为一个二维阵列 $f(x, y)$, 该阵列包含有 M 行 N 列, 其中 (x, y) 为离散坐标, 满足 $x = 0, 1, \dots, M - 1$ 与 $y = 0, 1, \dots, N - 1$ 。这样, 由一幅图像的坐标张成的实平面部分称为空间域, x 与 y 称为空间变量 (或空间坐标)。

- 我们也可以将这个 $M \times N$ 阵列表示成矩阵 $\mathbf{A} = (a_{ij})_{M \times N}$, 其中 $a_{ij} = f(x = i, y = j) = f(i, j)$ 。
- 【注】数字图像的原点 $(0, 0)$ 在左上角, 正 x 轴向下延伸, 正 y 轴向右延伸。
- 出于存储和量化硬件的考虑, 灰度级通常取 2 的正整数次幂, 即 $L = 2^k$; 当一幅图像有 2^k 个灰度级时, 通常称该图像为一幅 “ k 比特图像”。

- 存储数字图像需要的比特数 $b = MNk = MN \log_2 L$ 。
- 离散灰度级一般是等间隔的，同时是区间 $[0, L - 1]$ 内的整数，又被称为 **动态范围**，即图像系统中最大可度量灰度与最小可检测灰度之比。原则上，上线取决于**饱和度**，下限取决于**噪声**。
- **对比度**：一幅图像中最高和最低灰度级数之间的差定义为对比度。

空间分辨率：图像中可辨别的最小细节的度量，最通用的度量有以下两种。

- **单位距离的线对数**：假设线宽 W 个单位，那么线对宽度 $2W$ 单位，单位距离内有 $1/2W$ 个线对；
- **单位距离的点数（像素数）**：通常用每英寸像素数 (dpi) 表示。

灰度分辨率：在灰度级中可分辨的最小变化。

- **伪轮廓**：灰度级数不足时，数字图像平滑区域会出现轮廓。通常在 16 或更少级数的灰度图像中十分明显。

等偏爱曲线：改变数字图像的 N 与 k 并要求观察者主观对图像质量排序产生的 $N-k$ 平面等质量曲线。

图像内插：图像放大/收缩/旋转/几何校正中，需要用已知数据估测未知位置数据。

1. **最近邻内插**：将最靠近当前位置像素的灰度值赋给当前位置。简单，但是在某些直边缘会产生严重失真，有产生某些不希望的人为缺陷的倾向。
2. **双线性内插**： $v(x, y) = ax + by + cxy + d$ ，其中 a, b, c, d 由 4 个最近邻点灰度值写出的未知方程决定。效果比最近邻内插好得多，但是计算量增加。（注：双线性内插含交错项，不是线性内插）
3. **双三次内插**： $v(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j$ ，其中 16 个系数由 16 个最近邻点灰度值写出的未知方程决定。保持细节方面优于双线性内插。
4. 可以使用更多近邻点或更复杂的方法，如样条或小波。

2.5 像素间的一些基本关系

相邻像素

- 坐标 (x, y) 处的像素 p 有 4 个水平和垂直的相邻像素 $(x - 1, y)$, $(x + 1, y)$, $(x, y - 1)$ 与 $(x, y + 1)$ 。这组像素称为 p 的 **4 邻域**，用 $N_4(p)$ 表示；
- 坐标 (x, y) 处的像素 p 另外 4 个对角线相邻像素 $(x - 1, y - 1)$, $(x - 1, y + 1)$, $(x + 1, y - 1)$ 与 $(x + 1, y + 1)$ 。这组像素称为 p 的 **对角相邻像素**，用 $N_D(p)$ 表示；
- 4 个对角相邻像素与 4 个邻点一起称为 p 的 **8 邻域**，用 $N_8(p)$ 表示；
- 如果 (x, y) 位于图像的边界上，则 $N_4(p)$, $N_D(p)$ 和 $N_8(p)$ 的某些相邻像素位于数字图像的外部。

邻接性：令 V 表示用于定义邻接性的灰度值集合（包含当前考虑邻接性质的区域（点）），则

1. **4 邻接**：如果 $q \in N_4(p)$ ，那么具有 V 中数值的两个像素 p 和 q 是 4 邻接的；
2. **8 邻接**：如果 $q \in N_8(p)$ ，那么具有 V 中数值的两个像素 p 和 q 是 8 邻接的；
3. **m 邻接（混合邻接）**：如果 (i) $q \in N_4(p)$ 或 (ii) $q \in N_D(p)$ 且集合 $N_4(p) \cap N_4(q)$ 没有来自 V 中数值的像素，则具有 V 中数值的两个像素 p 和 q 是 m 邻接的。

在接下来的讨论中，必须提前给定邻接性的定义标准。

通路: 从具有坐标 (x, y) 的像素 p 到具有坐标 (s, t) 的像素 q 的通路 (或曲线) 是特定的像素序列, 其坐标为 $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$; 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}) 是邻接的 ($\forall i \in [1, n]$)。此时, n 为通路的长度。若 $(x_0, y_0) = (x_n, y_n)$, 则通路为闭合通路。根据邻接类型可分别定义 **4 通路**、**8 通路** 以及 m 通路。

连通分量: 令 S 是图像中的一个像素子集。我们说两个像素 p 与 q 在 S 中是联通的, 当且仅当 S 的全部像素之间存在一个 p 到 q 的通路。对于 S 中的任何像素 p , S 中联通到该像素的像素集称为 S 的一个连通分量。如果 S 仅有一个连通分量, 则其为**连通集**。

区域: 令 R 是图像中的一个像素子集。如果 R 是连通集, 则称 R 为一个区域。

- 两个区域 R_i, R_j 如果联合形成连通集, 则将 R_i 和 R_j 称为**邻接区域**;
- 不邻接的区域称为**不连接区域**。

前景与背景: 假设一幅图像包含 K 个不连接的区域, 即 R_1, R_2, \dots, R_K , 且他们都不接触图像的本身边界, 令 R_u 代表所有 K 个区域的并集, 则称 R_u 中所有点为图像的前景, $(R_u)^C$ 中所有点为图像的背景。

边界 (或边缘、轮廓) : 区域 R 的边界是这样的点集, 这些点和 R 的补集中的点临近; 也就是说边界为至少有一个背景邻点的像素集。

- 该定义也被称为“内边界”, 与背景的边界“外边界”区分。
- 如果 R 恰好是整幅图像, 那么边界由图像第一行、第一列、最后一行、最后一列的像素集合定义。

距离度量: 对于坐标分别为 (x, y) , (s, t) 和 (v, w) 的像素 p, q, z , 如果

1. $D(p, q) \geq 0$ ($D(p, q) = 0$ 当且仅当 $p = q$) ;
2. $D(p, q) = D(q, p)$;
3. $D(p, z) \leq D(p, q) + D(q, z)$ 。

则 D 是距离函数或距离度量。

一些特殊的距离度量

- 欧几里得 (欧氏) 距离: $D_e(p, q) = [(x - s)^2 + (y - t)^2]^{\frac{1}{2}}$ 。
- 曼哈顿距离 (城市街区距离) : $D_4(p, q) = |x - s| + |y - t|$
 $D_4 = 1$ 的像素是 (x, y) 的 4 邻域。
- 棋盘距离: $D_8(p, q) = \max(|x - s|, |y - t|)$
 $D_8 = 1$ 的像素是 (x, y) 的 8 邻域。
- 注意, D_4 距离与 D_8 距离与任何通路无关; 但是如果选择 m 邻接, 那么两个像素间的距离将依赖于其通路的像素值及其邻点值。

2.6 数字图像处理中所用数学工具的介绍 (略)

第三章 灰度变换与空间滤波

3.1 背景知识

本章讨论的空间域处理可以用下式表示:

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

其中, $f(x, y)$ 为输入图像, $g(x, y)$ 为处理后的图像, T 是在点 (x, y) 的邻域上定义的关于 f 的一种算子。

空间滤波: 邻域原点从一个像素向另一个像素移动, 对邻域中的像素应用 T 并在该位置产生输出, 然后输出 g 在这些坐标处的值, 该过程即为**空间滤波**; 其中邻域与与定义的操作算子 T 一起称为空间滤波器(或空间掩模、核、窗口、模板)。

灰度变换函数: 最小邻域大小为 1×1 , 那么 $g(x, y)$ 仅取决于 (x, y) 处的 f 值, 那么 T 可以看成一个单变量函数, 称为**灰度变换函数**

$$s = T(r)$$

这种结果仅取决于一个点灰度的方法, 称**点处理技术**; 结果取决于更大邻域的方法称**邻域处理技术**。

对比度拉伸: 如果 $T(r)$ 有图 7 左所示形式, 对 f 施以变换产生 g 将会让低于 k 的灰度级更暗, 高于 k 的灰度级更亮, 将比原是图像有更高的对比度, 因此称为**对比度拉伸**。**对比度拉伸**的极限情况为**阈值处理函数**, 如图 7 右所示, 将所有小于 k 的像素变到最暗, 将所有大于 k 的像素变到最亮。

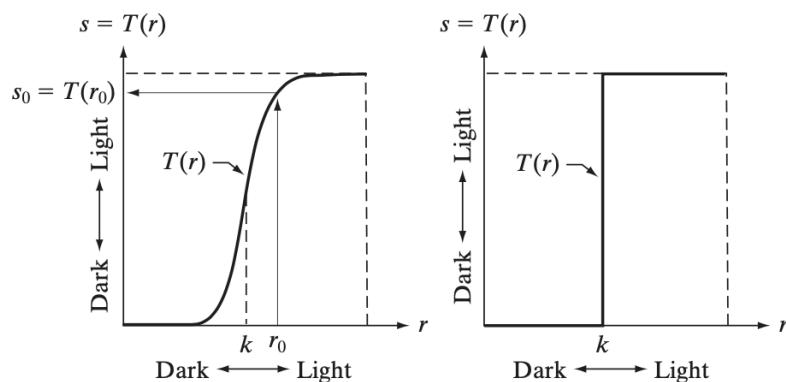


图 7

3.2 一些基本的灰度变换函数

图像反转: 反转图像的灰度级, 得到等效的照片底片。

$$s = L - 1 - r$$

在一幅图片中暗色区域占主导、不易观察时可通过反转使图片易于观察。

对数变换: 形如

$$s = c \log(1 + r)$$

的变换; 其中, c 是常数, 并且 $r \geq 0$, 如图 8 所示。

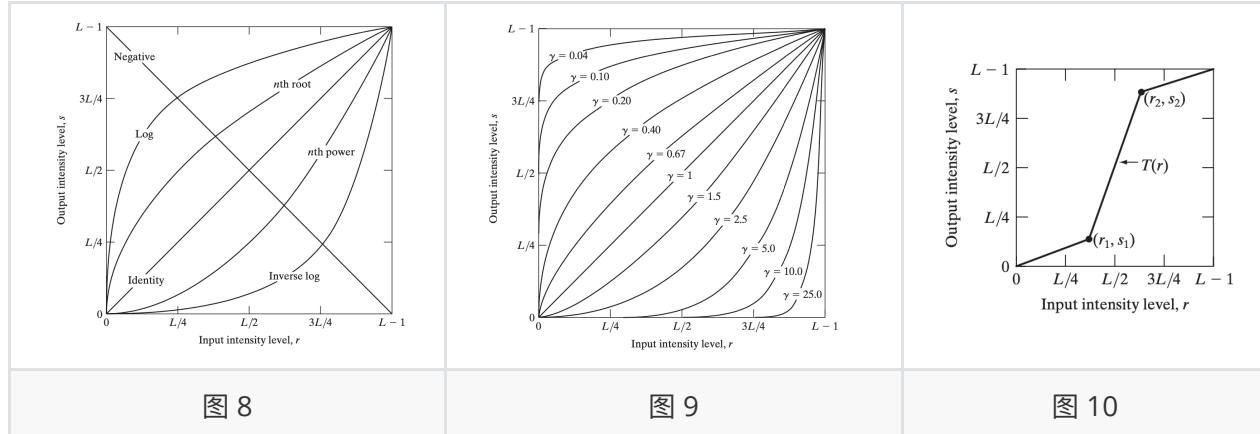
对数变换将输入中范围较窄的低灰度值映射为输出中范围较宽的灰度值, 获奖输入中范围较宽的高灰度值映射为输出中范围较窄的灰度值。我们通过这种类型的变换来扩展图像中的暗像素值、并且压缩更高灰度级的值。**反对数变换**的作用正好与之相反。

幂律(伽马)变换: 形如

$$s = cr^\gamma$$

的变换；其中， c 和 γ 为正常数，如图 9 所示；有时考虑偏移量（输入为 0 的一个可度量输出），改写为 $s = c(r + \varepsilon)^\gamma$ 。

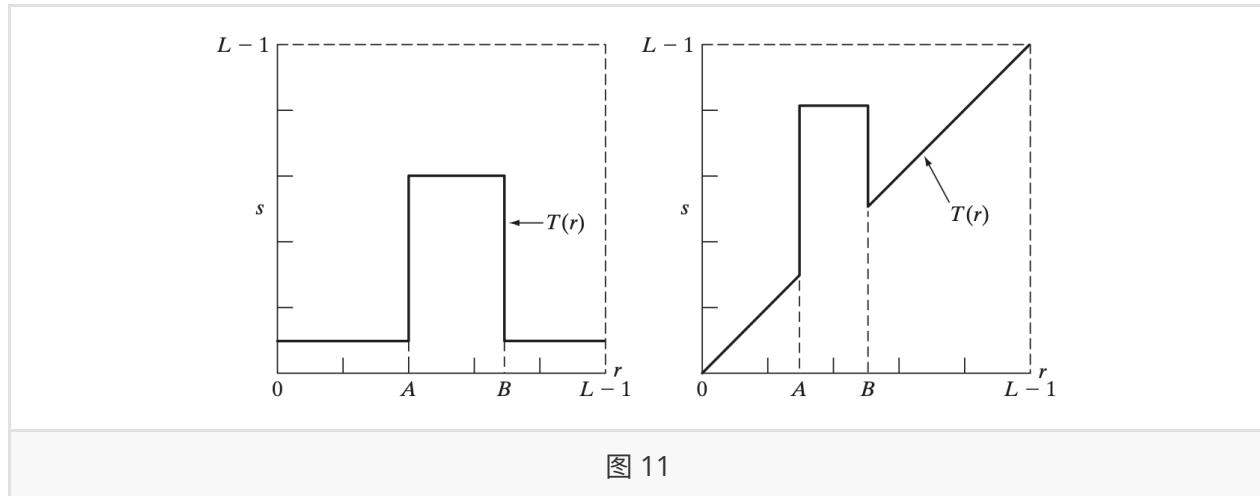
- $\gamma < 1$ 的伽马变换与对数变换的作用类似， γ 越小效果越明显； $\gamma > 1$ 的伽马变换与反对数变换的作用类似， γ 越大效果越明显。
- $c = \gamma = 1$ 时，伽马变换简化为恒等变换。
- 图像获取、打印和显示的各种设备通过幂律产生相应，因此用于校正这些幂律响应现象的处理称为 **伽马校正**；只需要在图像输入到监视器之前进行抵消设备幂律响应的伽马变换预处理即可。
- 幂律变换可以拉伸分辨率，增强细节。



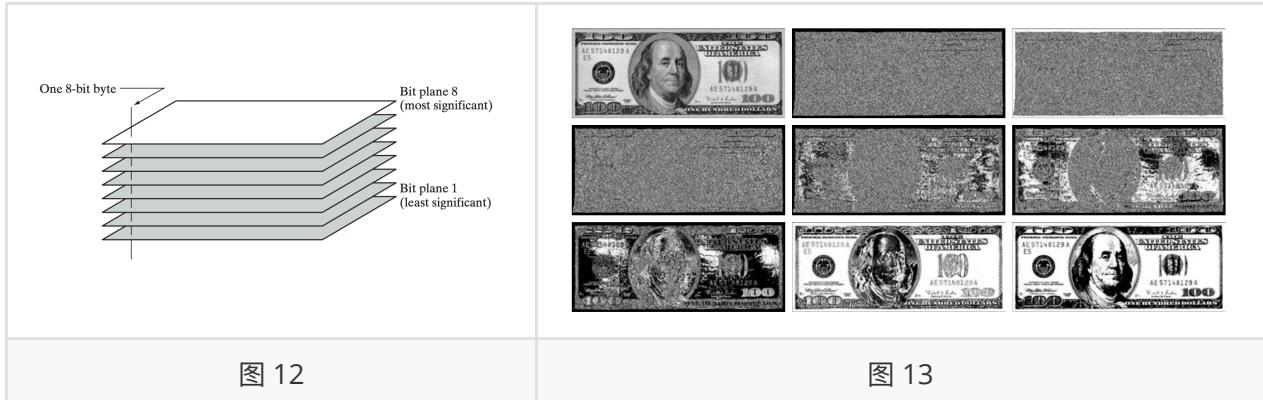
对比度拉伸：另一种对比度拉伸的函数如图 10 所示，其中点 (r_1, s_1) 、 (r_2, s_2) 的位置控制变换函数的形状。若 $r_1 = s_1$ 且 $r_2 = s_2$ 则变换为一线性恒等变换；若 $r_1 = r_2$ ， $s_1 = 0$ 且 $s_2 = L - 1$ ，则变换为阈值处理函数，产生一幅二值图像。

- 一般我们假设 $r_1 \leq r_2$ 且 $s_1 \leq s_2$ ，函数是单值的且单调递增，这样保证了灰度级的次序。
- 令 $(r_1, s_1) = (r_{min}, 0)$ ， $(r_2, s_2) = (r_{max}, L - 1)$ 可以进行简单的线性对比度拉伸，其中 $[r_{min}, r_{max}]$ 为原灰度级的范围。

灰度级分层：突出图像中特定灰度范围的亮度可以增强某些特征，通常称之为灰度级分层处理。主要有两大类方法：① 将感兴趣范围内的所有灰度值显示为一个值（如白色），而将其他灰度值显示为另一个值（黑色），如图 11 左所示；② 以线性恒等变换为基础，使感兴趣的灰度范围变亮，保持其他灰度不变，如图 11 右所示。



比特平面分层：一幅 8 比特图像由 8 个 1 比特平面构成，如图 12 所示；其中平面 1 包含图像中所有像素的最低阶比特，平面 8 包含图像中所有像素的最高阶比特。高阶比特在图像中贡献了视觉上很重要的大多数数据，低阶比特贡献了更精细的灰度细节，如图 14 所示。根据观察实验我们得出这样的结论：存储 4 个高阶比特平面将允许我们以可以接受的细节来重建原图；存储这 4 个平面代替原始图像可以减少 50% 的存储量（不考虑存储体系结构问题）。



3.3 直方图处理

数字图像的直方图：灰度级范围为 $[0, L - 1]$ 的数字图像的直方图是离散函数 $h(r_k) = n_k$ ，其中 r_k 是第 k 级灰度值， n_k 是图像中灰度为 r_k 的像素个数。实践中，经常用乘积 MN 表示的图像总像素除每个分量来归一化直方图，通常 M 和 N 为图像的行数和列数。归一化的直方图由

$$p(r_k) = \frac{n_k}{MN}$$

给出。简单地说， $p(r_k)$ 是灰度级 r_k 在图像中出现的概率的估计，归一化直方图的所有分量之和应该等于 1。

连续情况下的直方图均衡数学推导

假设连续灰度值情况下，令 r 表示待处理图像的灰度且 $r \in [0, L - 1]$ ，我们将注意力集中在变换形式上：

$$s = T(r) \quad 0 \leq r \leq L - 1$$

对于图像中每个具有 r 值的像素产生一个输出灰度值 s 。我们假设：

1. $T(r)$ 在区间 $[0, L - 1]$ 上单调递增；
2. $0 \leq r \leq L - 1$ 时， $0 \leq T(r) \leq L - 1$ 。

若要求 $T(r)$ 具有反函数 $r = T^{-1}(s)$ ，则将条件 1 改为“严格单调递增”。

根据概率论知识，我们有

$$p_s(s) = p_r(r) \left| \frac{dr}{ds} \right|$$

如果我们令

$$s = T(r) = (L - 1) \int_0^r p_r(w) dw$$

那么

$$\frac{ds}{dr} = (L - 1)p_r(r)$$

于是

$$p_s(s) = p_r(r) \cdot \frac{1}{(L - 1)p_r(r)} = \frac{1}{L - 1}$$

我们得到随机变量 s 是均匀分布的。

直方图均衡：离散情况下，我们进行如下变换

$$s_k = T(r_k) = (L - 1) \sum_{j=0}^k p_r(r_j) = \frac{L - 1}{MN} \sum_{j=0}^k n_j \quad (k = 0, 1, 2, \dots, L - 1)$$

则这种方式称为直方图均衡或直方图线性变化，该式可让灰度值的分布变得均衡。注意到直方图均衡是完全“自动的”，该式直接从已知图像提取信息，不需要更多的参数说明。

从 s 回到 r 的反变换形式为 $r = T^{-1}(s)$ ，只要原始图像中， r_k ($k \in [0, L - 1] \cap \mathbb{Z}$) 一个不缺，那么反变换就满足条件。

直方图匹配（直方图规定化）：要得到一幅灰度值分布为给定概率密度函数 $p_z(z_j)$ 的图像，如下操作：

1. 计算给定图像的直方图 $p_r(r)$ 并由上式得到 s_k ，把 s_k 四舍五入为范围 $[0, L - 1]$ 内的整数；
2. 令 $G(z_q) = (L - 1) \sum_{i=0}^q p_z(z_i)$ ，把 G 的值四舍五入为范围 $[0, L - 1]$ 内的整数并存入表中；
3. 把每个 s_k ($k = 0, 1, 2, \dots, L - 1$) 使用不走 2 存储的 G 值查询对应的 z_q 值，以使 $G(z_q)$ 最接近 s_k ，并储存这些 s 到 z 的映射。当满足给定 s_k 的 z_q 值多于 1 个时，选择最小的值；
4. 线对输入图像进行均衡得到 s_k ，然后使用步骤 3 存储的映射把图像中每个均衡后的像素值映射为直方图规定化后的图像中相应的 z_q 值，形成直方图规定化后的图像。

即现将原图进行直方图均衡，得到均衡图；在将目标直方图做直方图均衡，得到变换 $G(z_q)$ ，然后用反变换 $z_q = G^{-1}(s_k)$ 得到目标图。

局部直方图处理：定义一个邻域，并把该区域的中心从一个像素移到另一个像素；在每个位置，计算邻域中的点的直方图，并且在邻域内做直方图匹配/直方图规定化，得到对应变换函数并仅将其作用于区域的中心，得到中心对应的变换后的灰度值。

- 有时可以通过像素平移进行当前区域像素的增删，不必重新计算，加快计算速度；
- 可以使用菲重叠区域以减少计算量，但是可能产生“棋盘”效应。

图像统计变量：离散变量 r 关于其均值的 n 阶矩定义为

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

其中， m 是 r 的均值，即为图像平均灰度，即

$$m = \sum_{i=0}^{L-1} r_i p(r_i)$$

我们将二阶矩称为灰度方差，用 $\sigma^2 = \mu_2(r)$ 表示。均值 m 是平均灰度的度量，方差（或标准差 σ ）是图像对比度的度量。事实上，根据基本的统计学，均值和方差还可如下给出：

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

$$\sigma^2 = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (f(x, y) - m)^2$$

我们只需照上式计算，而不必计算直方图。图像局部的统计变量计算与全局基本相同，除了将计算范围限定在一个 (x, y) 为中心的邻域 S_{xy} 内，可以得出 $m_{S_{xy}}$ 与 $\sigma_{S_{xy}}^2$ 。

利用直方图统计进行图像增强：利用局部的均值 $m_{S_{xy}}$ 和方差 $\sigma_{S_{xy}}^2$ ，我们可以与全局均值 m_G 与方差 σ_G^2 比较，分别找出暗色区域以及低对比度区域，并将其做相应调整以增强图像。一个典型例子如下：

$$g(x, y) = \begin{cases} E \cdot f(x, y) & m_{S_{xy}} \leq k_0 m_G \text{ and } k_1 \sigma_G \leq \sigma_{S_{xy}} \leq k_2 \sigma_G \\ f(x, y) & \text{otherwise} \end{cases}$$

其中， E, k_0, k_1, k_2 为规定的参数。

3.4 空间滤波基础

线性空间滤波：使用大小为 $m \times n$ 的滤波器对大小为 $M \times N$ 的图像进行线性空间滤波，可以由下式表示：

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

其中， $m = 2a + 1, n = 2b + 1$ ；一般滤波器尺寸为奇数，也可以使用偶数尺寸或混合奇偶尺寸的滤波器，但是使用奇数滤波器简化索引、更为直观，并且让滤波器中心落在整数值上。有时，我们将线性空间滤波器写成向量表示 $R = \mathbf{w}^T \mathbf{z}$ ，其中 \mathbf{w} 为长度 $m \times n$ 的滤波器的系数展开而成的向量， \mathbf{z} 为滤波器覆盖的相应图像的灰度值。

空间相关：一个大小为 $m \times n$ 的滤波器 $w(x, y)$ 与一幅图像 $f(x, y)$ 做相关操作，可表示为 $w(x, y) \star f(x, y)$ ，由下式表示：

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

空间卷积：一个大小为 $m \times n$ 的滤波器 $w(x, y)$ 与一幅图像 $f(x, y)$ 做卷积操作，可表示为 $w(x, y) \star f(x, y)$ ，由下式表示：

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

离散单位冲激：将包含 1 个 1 和其余全部是 0 的函数称为离散单位冲激。

空间相关和卷积的讨论：在二维离散单位冲激中，空间相关的结果为滤波器翻转后的值；而空间卷积的结果为滤波器本身的值，如图 14 所示。在一些时候，如果没有歧义，我们直接使用 $g(x, y)$ 代替数字图像 $f(x, y)$ 与滤波器 $w(x, y)$ 卷积（或相关）的结果。

	Padded f	
Origin $f(x, y)$	0 0	
$w(x, y)$	0 1 2 3 0 0 0 0 0 0 0 4 5 6 0 0 0 0 0 0 0 7 8 9 0 0 0 0 0 0 0	
(a)	(b)	
Initial position for w	Full correlation result	Cropped correlation result
1 2 3 4 5 6 7 8 9	0 0	0 0 0 0 0 0 0 9 8 7 0 0 6 5 4 0 0 3 2 1 0 0 0 0 0 0
(c)	(d)	(e)
Rotated w	Full convolution result	Cropped convolution result
9 8 7 6 5 4 3 2 1	0 0	0 0 0 0 0 0 0 1 2 3 0 0 4 5 6 0 0 7 8 9 0 0 0 0 0 0
(f)	(g)	(h)

图 14

3.5 平滑空间滤波器

平滑线性滤波器（均值滤波器）： 平滑线性空间滤波器的输出是包含在滤波器模板邻域内的像素的简单平均值（或加权平均值），一般来说中间区域的权重会大于四周的权重。一幅 $M \times N$ 图像经过一个 $m \times n$ 点甲醛均值滤波器后的结果如下式给出：

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

- 所有系数都相等的空间均值滤波器称为**盒状滤波器**；
- 空间均值处理的一个重要应用是，模糊图像以便得到感兴趣物体的粗略描述（较小的物体的灰度将会和背景混合在一起，较大物体则变得易于检测）。

统计排序（非线性）滤波器： 统计排序滤波器是一种非线性空间滤波器，这种滤波器的响应以滤波器包围的图像区域中所包含的像素排序为基础，然后使用统计排序结果决定的值代替中心像素的值。

- 中值滤波器：** 用像素邻域内灰度的中值（在中值计算中包含原像素的值）代替该像素的值。种植滤波器提供了优秀的去噪能力，且比相同尺寸的线性平滑滤波器模糊程度明显要低。中值滤波器对去除脉冲噪声（又称为椒盐噪声，以黑白点叠加在图像上）非常有效。
- 最大值滤波器：** 用像素邻域内灰度的最大值代替该像素的值；
- 最小值滤波器：** 用像素邻域内灰度的最小值代替该像素的值。

3.6 锐化空间滤波器

锐化空间滤波器的目的是突出灰度的过渡部分，使图像变得更“尖锐”，细节更明显。

二阶微分的过零点对应图像边缘，因此二阶微分在增强细节方面比一阶微分好得多。

拉普拉斯算子： 一个二维图像 $f(x, y)$ 的拉普拉斯算子定义为

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

在 x 与 y 方向上，分别有

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

因此，两个变量的离散拉普拉斯算子为

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

使用拉普拉斯算子增强图像

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

事实上，拉普拉斯算子相当于如图 15 左上所示滤波器模板；可对算子做适当改变来得到图 15 右上、图 15 右下、图 15 左下的滤波器。

0	1	0	1	1	1
1	-4	1	1	-8	1
0	1	0	1	1	1
0	-1	0	-1	-1	-1
-1	4	-1	-1	8	-1
0	-1	0	-1	-1	-1

图 15 (左上为基本拉普拉斯算子，右上、右下、左下为基础变种)

非锐化掩蔽：

1. 模糊原图像；
2. 从原图像中减去模糊图像（产生的差值图像称为模板）；
3. 将模板加到原图像上。

令 $\bar{f}(x, y)$ 表示模糊图像，非锐化掩蔽以公式形式描述如下。首先，我们得到模板

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

然后，在原图像上加上该模板的一个权重部分

$$g(x, y) = f(x, y) + k \cdot g_{mask}(x, y)$$

通常，我们会在上式中包含一个权重系数 k ($k \geq 0$)；当 $k = 1$ 时，我们得到上面定义的非锐化掩蔽；当 $k > 1$ 时，该处理称为高提升滤波；选择 $k < 1$ 则不强调非锐化模板的贡献。

使用一阶微分（梯度）锐化（非线性）图像：对函数 $f(x, y)$ ， f 在坐标 (x, y) 处的梯度为二维列向量

$$\nabla f = \text{grad}(f) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \triangleq \begin{bmatrix} g_x \\ g_y \end{bmatrix}$$

向量 ∇f 的幅值表示为 $M(x, y)$, 即 $M(x, y) = \text{mag}(\nabla f) = \sqrt{g_x^2 + g_y^2}$ 。

有时, 也用绝对值来近似平方和, 即 $M(x, y) \approx |g_x| + |g_y|$ 。

在离散情况下, 我们有罗伯特交叉梯度算子, 即为

$$M(x, y) = [(z_9 - z_5)^2 + (z_8 - z_6)^2]^{\frac{1}{2}} \approx |z_9 - z_5| + |z_8 - z_6|$$

其中, z_1, z_2, \dots, z_9 的定义如图 16 所示, 罗伯特交叉梯度算子的模板如图 16 所示。由于罗伯特交叉梯度算子的尺寸为偶数, 难以实现; 我们用另一种方法进行对 g_x 和 g_y 的近似, 即

$$g_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3); \quad g_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

从而得到 **Sobel 算子**, 即

$$M(x, y) \approx |(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)| + |(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)|$$

Sobel 算子的模板如图 16 所示。

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

-1	0	0
0	1	-1
1	0	0

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

图 16 (最上方为 z 的定义, 中间为两个罗伯特交叉梯度算子, 最后一行为两个Sobel算子)

使用这些模板可以得到各个位置对应的 g_x 与 g_y , 使用任意一种方法 (近似) 可以得到梯度模长 $M(x, y)$, 从而进行得到对应的边缘细节信息, 达到锐化的目的 (一般用来增强边缘)。

3.7 混合空间增强法

通过对上述方法进行组合、灵活运用来增强图像。

3.8 用模糊技术进行灰度变换和空间滤波 (略)