

数字图像处理

(Digital Image Processing)

核心知识点精炼笔记 1-2

东南大学计算机科学与工程学院
09023419 严伟琛

授课教师：薛澄
参考教材：拉斐尔·冈萨雷斯 (Rafael C. Gonzalez) 第四版
2026 年 1 月 15 日

目录

1 緒論 (Introduction)	2
1.1 图像的基本定义与意义	2
1.2 数字图像的本质与数学描述	2
1.3 图像工程的三个层次	2
1.4 学科关系辨析	3
1.5 应用领域: 按信息源分类	3
1.6 数字图像处理的基本步骤	3
1.7 系统硬件组成与典型示例 (CT)	4
1.8 第一章核心重点问题总结	4
2 数字图像处理基础知识 (Digital Image Fundamentals)	5
2.1 人类视觉系统	5
2.1.1 眼球结构与感光细胞	5
2.1.2 亮度适应与辨别	5
2.1.3 常见的视觉现象	5
2.2 图像感知和获取	6
2.3 图像取样和量化	6
2.3.1 例题 1: 空间分辨率单位换算	6
2.3.2 例题 2: CCD 成像芯片选型与面积计算	7
2.3.3 例题 3: 数据包协议下的传输时间计算	7
2.4 像素间的基本关系	8
2.4.1 像素的邻域	8
2.4.2 邻接性、连通性与通路	8
2.4.3 通路与连通性	8
2.4.4 例题 4: 通路长度计算	9
2.4.5 例题 5: 图像子集的邻接性判定	10
2.4.6 距离度量	11
2.4.7 例题 6: 像素间距离计算	12
2.5 基本数学工具	13
2.5.1 线性算子与非线性算子	13
2.5.2 算术运算及其典型应用	13
2.5.3 几何空间变换 (Affine Transformations)	13
2.5.4 灰度内插	14
2.5.5 例题 7: 通用仿射变换矩阵的逆变换	15
2.6 第二章核心重点问题总结	16

1 绪论 (Introduction)

1.1 图像的基本定义与意义

- **图像定义**: 图像是对客观存在的物体的一种相似的、生动的写真或描述。
- **视觉的重要性**: 人类从外界获得的信息约有 **60%~90%** 来自视觉系统。作为传递信息的媒介，图像具有极高的地位。
- **信息带宽对比** (潜在考点):
 - 视觉: 1250 MB/S
 - 听觉/嗅觉: 125 MB/S
 - 触觉: 12.5 MB/S

1.2 数字图像的本质与数学描述

- **连续图像**: 定义为二维连续函数 $f(x, y)$, 其中 (x, y) 是空间坐标, f 代表该点物理性质 (如亮度) 的数值。
- **数字图像**: 二维离散函数。其空间坐标 (x, y) 和幅值 $f(x, y)$ 均经过了**离散化** (取样与量化)。
- **像素 (Pixel)**: 由离散的位置和灰度值组成, 是数字图像的基本单元。
- **常见格式**:
 - 单属性: 黑白或灰度图像。
 - 三属性: RGB (红、绿、蓝)。
 - 四属性: RGBA (增加透明度通道)。

1.3 图像工程的三个层次

图像工程根据处理对象的抽象程度, 分为三个紧密相关的层次:

1. **低层处理 (图像处理)**:
 - **输入/输出**: 均是图像。
 - **典型操作**: 对比度增强、图像平滑 (去噪)、图像锐化。
 - **操作对象**: 像素。
2. **中层处理 (图像分析)**:
 - **输入**: 图像; **输出**: 图像属性 (特征)。
 - **典型操作**: 图像分割、物体识别、区域特征描述。
 - **操作对象**: 目标物体。
3. **高层处理 (图像理解)**:
 - **目标**: 理解图像内容, 模拟人类视觉, 感知三维场景。
 - **操作对象**: 符号/语义。

规律: 从低层到高层, 抽象程度越高, 而数据量越小。

1.4 学科关系辨析

- **与数字信号处理 (DSP) 的关系:** DSP 处理一维信号, 图像处理涉及二维及以上信号。两者在滤波器设计、正交变换、编码等方面紧密相关。
- **与计算机图形学 (CG) 的关系:** CG 是“由简至繁”, 通过数学公式生成仿真图像; 图像处理则是“由繁至简”, 从原始图像中提取结果或进行变换。
- **与计算机视觉 (CV) 的关系:** CV 侧重于利用特征感知、识别并理解三维场景, 属于高层处理范畴。

1.5 应用领域: 按信息源分类

- **电磁波谱成像:**
 - **伽马射线:** 核医学 (PET)、天文观测。
 - **X 射线:** 医学 (CT、胸透)、工业探伤、昆虫化石科研。
 - **紫外线:** 荧光显微镜显微、工业检测。
 - **可见光/红外:** 遥感 (卫星监测)、气象、指纹/人脸识别、显微镜成像。
 - **微波:** 雷达 (星载雷达图像)。
 - **无线电波:** 核磁共振 (MRI)。
- **非电磁波成像:**
 - **声波:** B 超 (医学)、油气阱地质勘探 (低频声波)。
 - **电子:** 扫描电子显微镜 (SEM)、透射电子显微镜 (TEM)。
- **计算机生成图像:** 分形算法、三维模型仿真。

1.6 数字图像处理的基本步骤

一个完整的处理系统逻辑上包含以下模块:

- **图像获取:** 数字化与预处理。
- **图像增强:** 改善质量。特点: 具有主观性, 为了使图像更适用于特定应用。
- **图像复原:** 修复退化。特点: 具有客观性, 以图像退化的数学模型为基础。
- **小波/多分辨率处理:** 不同分辨率描述, 实现数据压缩。
- **压缩:** 减少存储量或传输带宽 (JPEG, GIF, PNG)。
- **形态学处理:** 提取描述形状的图像分量。
- **分割与识别:** 划分目标并赋予标记 (字符识别、指纹/人脸识别)。
- **知识库 (Knowledge Base):** 蕴含问题的先验知识 (如感兴趣区位置), 与各模块双向交互。

1.7 系统硬件组成与典型示例 (CT)

通用图像处理系统主要构成：

- **图像传感器**: 获取原始信息 (摄像机、显微镜)。
- **专用硬件**: 数字化仪、算术逻辑单元 (ALU, 常用于求平均去噪)。
- **大容量存储**:
 - **短期存储**: 内存、帧缓存。
 - **在线存储**: 硬盘。
 - **档案存储**: 磁带、光盘。

典型系统示例: CT (计算机断层成像):

- **核心元器件**: 球管、高压发生器、滑环、探测器。
- **基本过程**: 数据采集 (正投影) → 投影校正 → 图像重建 → 三维渲染。
- **数学本质**: 利用线积分 (Radon 变换) 描述:

$$p_\theta(s) = \iint f(x, y) \delta(s - x \cos \theta - y \sin \theta) dx dy \quad (1)$$

1.8 第一章核心重点问题总结

1. **什么是数字图像处理**? 借助数字计算机, 通过算法对经过离散化 (取样与量化) 处理的二维/多维信号进行加工, 以改善图像质量或提取有用信息的过程。
2. **数字图像处理的起源是什么**? 起源于 20 世纪 20 年代的报纸图像电缆传输 (Bartlane 系统), 在 60 年代因航天技术 (如旅行者号探测器) 和数字计算机的发展而成熟, 并在 70 年代随着 CT 技术的发明进入医疗领域。
3. **数字图像处理的应用领域有哪些**? 涵盖了全电磁波谱波段 (从高能量的伽马射线到低能量的无线电波), 以及非电磁波成像 (如超声波声导成像、电子显微镜成像) 和计算机生成图像。
4. **数字图像处理的基本步骤有哪些**? 完整的流水线包括: 图像获取 → 滤波与增强 → 复原 → 彩色处理 → 小波/多分辨率分析 → 压缩 → 形态学处理 → 分割 → 特征提取 → 目标识别。其中知识库是核心支撑。
5. **图像处理系统的组成要素有哪些**? 由硬件 (图像传感器、专用处理硬件、计算机、大容量存储、显示器、网络) 和软件 (特定任务模块、可编程调用接口) 共同构成。

2 数字图像处理基础知识 (Digital Image Fundamentals)

2.1 人类视觉系统

研究人类视觉系统 (HVS) 的目的在于：

- **主观评价标准**: 数字图像的质量最终多由人眼进行主观评价。
- **技术设计依据**: 许多技术 (如屏幕刷新频率、图像量化级别、图像压缩算法) 均是根据人眼生理特性设计的。
- **认知机制参考**: 理解人类如何感知图像内容，对于开发自动识别算法至关重要。

2.1.1 眼球结构与感光细胞

人眼构造类似于相机，光线通过角膜、晶状体聚焦在视网膜上。视网膜分布着两类关键的感光细胞：

1. 锥状细胞：

- **数量**: 约 600 万–700 万个。
- **分布**: 主要集中在视网膜中心部位的中央凹。
- **功能**: 对颜色高度敏感，空间分辨率高，负责“明视觉” (Photopic vision)。

2. 杆状细胞：

- **数量**: 约 7500 万–1.5 亿个。
- **分布**: 分布在视网膜表面。
- **功能**: 没有色彩感觉，空间分辨率低，但在低照明条件下极其灵敏，负责“暗视觉” (Scotopic vision)。

2.1.2 亮度适应与辨别

- **主观亮度**: 人眼感知到的亮度 (灰度) 并不是物理强度的线性函数，而是一个对数函数。
- **亮度适应级别**: 人眼不能同时辨别整个亮度范围，而是通过改变其视觉灵敏度级别来适应背景亮度。
- **韦伯比 (Weber Ratio)**: 定义为 $\frac{\Delta I_c}{I}$ 。
 - I 是背景强度， ΔI_c 是人眼能辨别出的最小亮度变化。
 - **结论**: 韦伯比较小时，表示人眼对亮度变化更敏感 (辨别能力强)；在低照明水平下，韦伯比通常较大。

2.1.3 常见的视觉现象

- **马赫带现象 (Mach Bands)**: 在图像中不同强度区域的边界处，人眼会感知到“下冲”或“上冲”的错觉。这种现象证明了人类视觉系统具有**锐化边缘**的特性，使得边界看起来比实际物理强度对比更明显。

- **同时对比 (Simultaneous Contrast)**: 同一个灰度块在白色背景下看起来比在黑色背景下更暗。这说明感知亮度不仅取决于该点本身，还取决于其背景环境。
- **视觉惰性**: 光消失后，人眼的感觉会持续一段时间。
 - **临界闪烁频率**: 当光脉冲频率增加到某一值时，人眼感觉不到闪烁。
 - **数值**: 人眼的临界闪烁频率约为 **46Hz**。

2.2 图像感知和获取

图像的获取通常涉及一个照射源和对该源敏感的传感器。

- **基本过程**:
 - 光电转换**: 传感器接收目标反射或透射的辐射能量，产生与照射能量强度成正比的模拟电信号。
 - 模数转换 (ADC)**: 将模拟电信号转化为数字信号，以便计算机处理。
- **传感器类型**:
 - **单个传感器**: 通过传感器在二维平面内的机械运动（如左右、旋转）来生成图像。
 - **条带传感器**: 一维排列。常见于扫描仪，或通过物体运动进行感知（如 CT 的环形条带）。
 - **阵列传感器**: 如 CCD 或 CMOS。传感器呈二维排列，可以一次性获取整个图像平面的信息。

2.3 图像取样和量化

要将连续图像 $f(x, y)$ 转换为数字形式，必须在空间坐标和幅值两个维度上进行离散化处理。

- **取样**: 对空间坐标 (x, y) 的离散化。取样频率直接决定了图像的**空间分辨率**。取样越细，图像细节表现越清晰；若取样不足，则会出现“棋盘模式”或锯齿状边缘。
- **量化**: 对函数幅值（灰度值）的离散化。量化层级 L 决定了图像的**灰度分辨率**。出于存储和硬件考虑，灰度级通常取为 2 的整数次幂，即 $L = 2^k$ （常见为 8 比特，即 256 级）。量化级数过低会导致图像出现“虚假轮廓”现象。
- **数字表示与存储**: 一幅大小为 $M \times N$ 像素、灰度级为 $L = 2^k$ 的图像所需的存储空间计算公式为：

$$b = M \times N \times k \quad (\text{单位: bit}) \quad (2)$$

2.3.1 例题 1：空间分辨率单位换算

题目: 假设你必须将一幅大小为 2048×2048 像素的图像插入报告中。

- 要使图像打印在 $5\text{cm} \times 5\text{cm}$ 的空间上，分辨率必须是多少线对每毫米 (lp/mm)?
- 要使图像打印在 2 英寸 \times 2 英寸的空间上，分辨率必须是多少 dpi?

详细解答步骤:

1. 解答 (a):

- 第一步：单位换算。将厘米转换为毫米： $5\text{cm} = 50\text{mm}$ 。
- 第二步：理解线对概念。1个线对 (lp) 由1条黑线和1条白线组成，至少需要2个像素来表示。
- 第三步：应用公式。分辨率 $(\text{lp/mm}) = \frac{\text{单边像素数}}{\text{打印尺寸 (mm)} \times 2}$ 。
- 计算： $\frac{2048}{50 \times 2} = \frac{2048}{100} = 20.48 \text{ lp/mm}$ 。

2. 解答 (b):

- 第一步：理解 dpi 含义。dpi 指每英寸点（像素）数。
- 第二步：应用公式。分辨率 $(\text{dpi}) = \frac{\text{单边像素数}}{\text{打印尺寸 (英寸)}}$ 。
- 计算： $\frac{2048}{2} = 1024 \text{ dpi}$ 。

2.3.2 例题 2：CCD 成像芯片选型与面积计算

题目：某个自动成像应用需要 5 lp/mm 的最小分辨率。摄像机焦点与被摄区域距离为 1m ，成像面积为 $0.5\text{m} \times 0.5\text{m}$ 。配备 200mm 的镜头。满足要求的 CCD 芯片的最小感测元素的数量和面积 ($d \times d$) 是多少？

详细解答步骤：

1. 第一步：计算成像尺寸。利用相似三角形原理，像高 h_i 与物高 h_o 之比等于像距（焦距） d_i 与物距 d_o 之比。

$$h_i = h_o \times \frac{d_i}{d_o} = 500\text{mm} \times \frac{200\text{mm}}{1000\text{mm}} = 100\text{mm} \quad (3)$$

2. 第二步：确定所需像素数量。分辨率要求 5 lp/mm ，意味着每毫米需要 $5 \times 2 = 10$ 个像素。

$$\text{单边像素数} = 100\text{mm} \times 10 \text{ 像素/mm} = 1000 \text{ 像素} \quad (4)$$

因此，最小感测元素数量为 $1000 \times 1000 = 1.0 \times 10^6$ (100 万像素)。

3. 第三步：计算感测元素面积。

- 像素边长 $d = \frac{\text{成像总高度}}{\text{像素数}} = \frac{100\text{mm}}{1000} = 0.1\text{mm}$ 。
- 像素面积 $= d \times d = 0.1\text{mm} \times 0.1\text{mm} = 0.01 \text{ mm}^2$ 。

2.3.3 例题 3：数据包协议下的传输时间计算

题目：传输是以数据包形式实现的，每个数据包含 1 个开始位、1 个字节（8 比特）信息和 1 个结束位。传输 500 幅大小为 1024×1024 的 256 灰度级图像，分别计算以下波特率所需的时间：
(a) 3M 波特 ($3 \times 10^6 \text{ bit/s}$); (b) 30G 波特 ($30 \times 10^9 \text{ bit/s}$)。

详细解答步骤：

1. 第一步：计算单像素实际传输比特数。256 灰度级需要 8 比特。由于每个字节外加 1 开始位和 1 结束位，实际每像素传输比特为 $8 + 1 + 1 = 10 \text{ bit}$ 。
2. 第二步：计算总数据量。

$$\text{总比特数} = 1024 \times 1024 \times 10 \text{ (bit/pixel)} \times 500 \text{ (幅)} = 5,242,880,000 \text{ bit} \quad (5)$$

3. 第三步：代入公式计算时间（时间 = 总数据量 / 波特率）。

- (a) 3M 波特： $\frac{5.24288 \times 10^9}{3 \times 10^6} \approx 1747.63$ 秒。
- (b) 30G 波特： $\frac{5.24288 \times 10^9}{30 \times 10^9} \approx 0.17476$ 秒。

2.4 像素间的基本关系

2.4.1 像素的邻域

对于坐标为 (x, y) 的像素 p , 其常见的邻域定义如下:

- 4 邻域 $N_4(p)$: 包含 p 的 4 个垂直和水平相邻像素, 坐标为 $(x \pm 1, y)$ 和 $(x, y \pm 1)$ 。
- D 邻域 $N_D(p)$: 包含 p 的 4 个对角相邻像素, 坐标为 $(x \pm 1, y \pm 1)$ 。
- 8 邻域 $N_8(p)$: 4 邻域与 D 邻域的并集, $N_8(p) = N_4(p) \cup N_D(p)$ 。

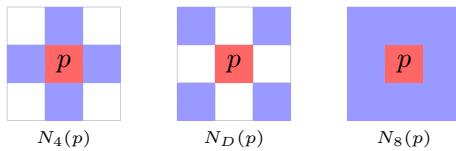


图 1: 像素邻域示意图: 红色为中心像素 p , 蓝色为其相邻像素

提示: 位于图像边缘的像素, 其邻域可能会超出图像边界。

2.4.2 邻接性、连通性与通路

为了定义两个像素是否“接触”, 必须引入灰度值集合 V (相似性准则)。例如, 在二值图像中, 通常令 $V = \{1\}$ 。

1. 4 邻接: 如果 $q \in N_4(p)$, 且 p, q 的灰度值均属于 V 。
2. 8 邻接: 如果 $q \in N_8(p)$, 且 p, q 的灰度值均属于 V 。
3. m 邻接 (混合邻接): 满足以下任一条件即为 m 邻接:

- $q \in N_4(p)$, 且 p, q 的灰度值属于 V 。
- $q \in N_D(p)$, 且 p, q 的灰度值属于 V , 并且 $N_4(p) \cap N_4(q)$ 的交集中没有任何像素的灰度值属于 V 。

核心意义: m 邻接是为了消除使用 8 邻接时常出现的“多路径二义性”。

2.4.3 通路与连通性

- **通路 (Path):** 从像素 p 到 q 的一个像素序列, 序列中相邻的像素均满足特定的邻接性 (4、8 或 m 邻接)。
- **连通性:** 若 p, q 之间存在通路, 则称它们是连通的。根据通路类型可分为 4 连通、8 连通和 m 连通。

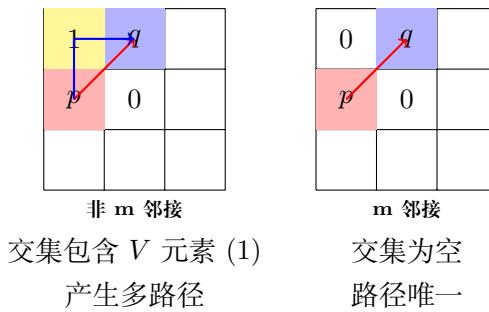


图 2: m 邻接判定: 当 $V = \{1\}$ 时, 左图因存在中转点 (1) 而不构成 m 邻接, 右图则构成。

典型应用示例: 在一幅包含 0, 1, 2 灰度值的图像中, 如果 $V = \{0, 1\}$, 从左下角 p 到右上角 q :

1. **4 通路:** 可能不存在。因为 q 点的垂直和水平邻域内可能没有属于 V 的像素。
2. **8 通路:** 通常最短, 因为允许对角线跳跃 [3]。
3. **m 通路:** 长度通常介于 4 和 8 之间, 其目的是确保提取出的边界是单像素宽的。

2.4.4 例题 4: 通路长度计算

题目: 考虑下图所示的图像分割。令 p 为左下角坐标为 $(0, 0)$ 的像素, q 为右上角坐标为 $(3, 3)$ 的像素。

- (a) 令 $V = \{0, 1\}$ 是定义邻接性的灰度值集合。请计算 p 与 q 之间最短的 4-通路、8-通路和 m-通路的长度。若其中某种通路不存在, 请说明原因 [1]。
- (b) 令 $V = \{1, 2\}$, 重做 (a) 问 [1]。

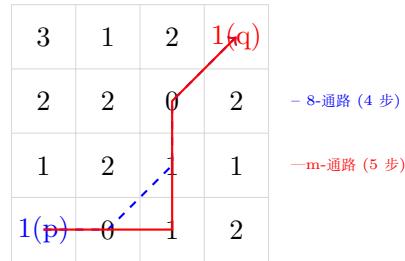
	3	1	2	1(q)
2	2	2	0	2
1	1	2	1	1
0	1(p)	0	1	2

图 3: 通路计算原始网格

详细解答步骤:

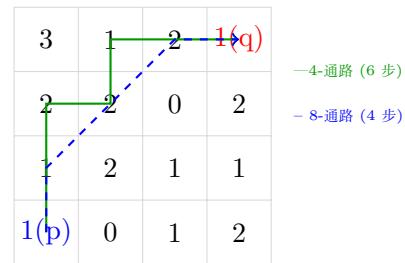
- (a) 当相似性准则集合 $V = \{0, 1\}$ 时:

- **4-通路:** 不存在。原因: 终点 $q(3, 3)$ 的 4-邻域内像素值为 2 (左) 和 2 (下), 均不属于 $V = \{0, 1\}$, 导致通路中断。
- **8-通路序列 (最短长度 4):** $(0, 0) \xrightarrow{\text{右}} (1, 0) \xrightarrow{\text{右上}} (2, 1) \xrightarrow{\text{上}} (2, 2) \xrightarrow{\text{右上}} (3, 3)$ 。
- **m-通路序列 (最短长度 5):** 由于 $(1, 0)$ 与 $(2, 1)$ 的公共 4-邻域中点 $(2, 0) = 1 \in V$, 按 m -邻接规则必须绕行: $(0, 0) \xrightarrow{\text{右}} (1, 0) \xrightarrow{\text{右}} (2, 0) \xrightarrow{\text{上}} (2, 1) \xrightarrow{\text{上}} (2, 2) \xrightarrow{\text{右上}} (3, 3)$ 。

图 4: $V = \{0, 1\}$ 时的完整路径标注

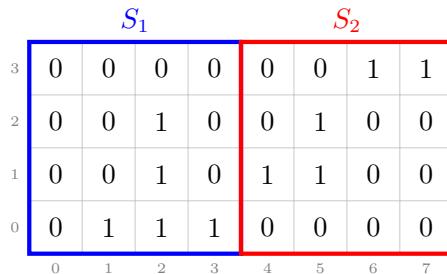
(b) 当相似性准则集合 $V = \{1, 2\}$ 时:

- 最短 4-通路 (长度 6): $(0,0) \xrightarrow{\text{上}} (0,1) \xrightarrow{\text{上}} (0,2) \xrightarrow{\text{右}} (1,2) \xrightarrow{\text{上}} (1,3) \xrightarrow{\text{右}} (2,3) \xrightarrow{\text{右}} (3,3)$ 。
- 最短 8-通路 (长度 4): $(0,0) \xrightarrow{\text{上}} (0,1) \xrightarrow{\text{右上}} (1,2) \xrightarrow{\text{右上}} (2,3) \xrightarrow{\text{右}} (3,3)$ 。
- 最短 m-通路 (长度 6): 由于 8-通路中的对角跳跃 (如 $(0,1) \rightarrow (1,2)$) 其公共邻域 $(0,2) = 2 \in V$, 根据 m -邻接定义, 这些对角路径均需断开, 因此其最短路径与 4-通路重合, 长度为 6 [2]。

图 5: $V = \{1, 2\}$ 时的完整路径标注

2.4.5 例题 5: 图像子集的邻接性判定

题目: 已知两个图像子集 S_1 和 S_2 如下图所示。假设相似性准则灰度集合为 $V = \{1\}$, 请确定这两个子集之间是否满足: (a) 4-邻接; (b) 8-邻接; (c) m -邻接。若不满足, 请简述理由。

图 6: 子集 S_1 与 S_2 的原始灰度分布网格

详细解答步骤:

1. 4-邻接判定:

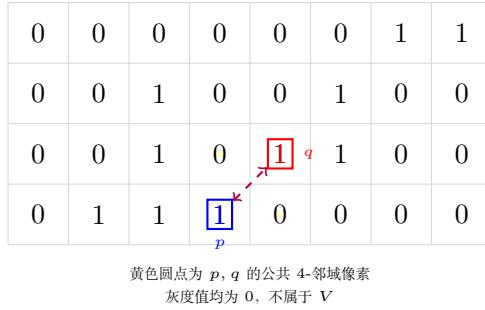


图 7: 邻接性分析图示（紫色虚线表示关键邻接点对）

- 检查 S_1 中的所有 “1” 像素与 S_2 中的所有 “1” 像素。
- 若要满足 4-邻接，必须存在至少一对点 (p, q) 使得 $q \in N_4(p)$ (即水平或垂直相邻)。
- 观察边界: S_1 在第 3 列的 “1” 位于 $(3, 0)$, 而 S_2 在第 4 列的 “1” 位于 $(4, 1)$ 。它们处于对角线位置。
- 经遍历，不存在任何一对水平或垂直相邻的 “1” 像素点对。
- 结论: S_1 与 S_2 不是 4-邻接的。

2. 8-邻接判定:

- 检查是否存在 $q \in N_8(p)$ 。
- 像素 $p(3, 0) \in S_1$ 与像素 $q(4, 1) \in S_2$ 均为 “1”，且它们在地理上是对角相邻的。
- 只要存在至少一对满足条件的像素，子集即满足邻接性 [1]。
- 结论: S_1 与 S_2 是 8-邻接的。

3. m-邻接判定:

- 根据 m-邻接定义，对于对角相邻的点对 (p, q) ，需检查其公共 4-邻域 $N_4(p) \cap N_4(q)$ 。
- 点 $p(3, 0)$ 和 $q(4, 1)$ 的公共 4-邻域点为 $(3, 1)$ 和 $(4, 0)$ 。
- 查表可知，这两个点在图像中的灰度值均为 0。
- 因为 $0 \notin V$ ，即公共邻域中没有属于 V 的元素，满足 m-邻接的对角连接条件 [1]。
- 结论: S_1 与 S_2 是 m-邻接的。

2.4.6 距离度量

对于坐标分别为 $p(x, y)$ 、 $q(u, v)$ 和 $z(s, t)$ 的像素，一个有效的距离函数 D 必须满足三个基本性质：非负性 ($D(p, q) \geq 0$, 当且仅当 $p = q$ 时为 0)、对称性 ($D(p, q) = D(q, p)$) 以及三角不等式 ($D(p, z) \leq D(p, q) + D(q, z)$)。

常用距离度量定义及其几何轮廓:

1. 欧氏距离 (Euclidean Distance, D_e): 定义为两点之间的直线距离。

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

几何特征: 与点 (x, y) 的欧氏距离小于等于某个值 r 的所有像素组成一个以该点为圆心的圆平面。

2. 城市距离 (City-block Distance, D_4): 又称曼哈顿距离。定义为像素只能沿水平和垂直方向移动时的最短路径。

$$D_4(p, q) = |x - u| + |y - v|$$

几何特征: 与点 (x, y) 的 D_4 距离小于等于 r 的所有像素组成一个菱形。具有 $D_4 = 1$ 的像素集合即为该点的 4-邻域 (N_4)。

3. 棋盘距离 (Chessboard Distance, D_8): 定义为像素在行或列上移动的最大位移量，模拟国际象棋中“王”的走法。

$$D_8(p, q) = \max(|x - u|, |y - v|)$$

几何特征: 与点 (x, y) 的 D_8 距离小于等于 r 的所有像素组成一个正方形。具有 $D_8 = 1$ 的像素集合即为该点的 8-邻域 (N_8)。



图 8: 欧氏距离 D_e

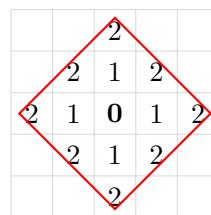


图 9: 城市距离 D_4

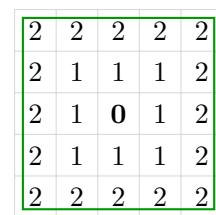


图 10: 棋盘距离 D_8

2.4.7 例题 6: 像素间距离计算

题目: 已知图像中两个像素点 p 和 q 的坐标分别为 $(1, 1)$ 和 $(6, 3)$ 。请计算它们之间的欧氏距离 (D_e)、城市距离 (D_4) 和棋盘距离 (D_8)。

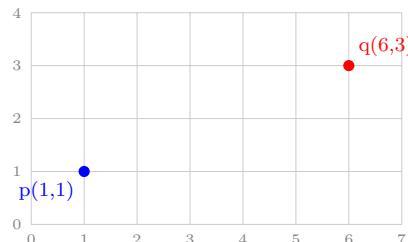


图 11: 像素点 p 与 q 的位置分布

详细解答步骤:

1. 第一步: 计算坐标差绝对值:

- 水平跨度: $\Delta x = |6 - 1| = 5$
- 垂直跨度: $\Delta y = |3 - 1| = 2$

2. 第二步: 代入公式计算:

- 欧氏距离 (D_e): $D_e = \sqrt{5^2 + 2^2} = \sqrt{29} \approx 5.39$ 。
- 城市距离 (D_4): $D_4 = 5 + 2 = 7$ 。
- 棋盘距离 (D_8): $D_8 = \max(5, 2) = 5$ 。

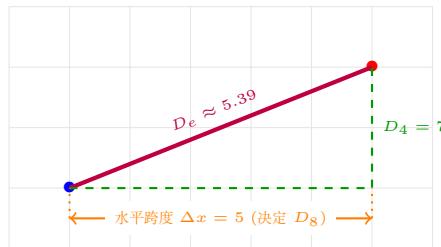


图 12: 三种距离的几何意义对比

2.5 基本数学工具

数字图像处理的实现依赖于多种数学工具。根据理论体系，这些工具涵盖了从基础算术运算到复杂的空间几何变换和灰度内插算法。

2.5.1 线性算子与非线性算子

- **线性算子**: 算子 H 必须同时满足加性 (additivity) 和同质性 (homogeneity)。即对于任意两幅图像 f_1, f_2 及常数 a, b , 均有:

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

常见的加法、减法、求平均等操作均为线性运算。

- **非线性算子**: 不满足线性叠加性质的操作。典型示例包括图像处理中的最大值 (Max)、最小值 (Min) 操作以及中值滤波 (Median)。

2.5.2 算术运算及其典型应用

图像间的算术运算是像素级对对应位置的元素执行操作:

- **加法运算 (去噪)**: 通过对 K 幅含有零均值、独立噪声的图像求平均, 可以有效抑制噪声。
- **减法运算 (差异增强)**: 用于突出两幅图像之间的变化。在医疗领域中常用于数字减影血管造影 (DSA), 通过减去背景图来显现显影后的血管。
- **乘法/除法运算**: 常用于校正阴影 (光照不均的影响) 或通过模板 (Mask) 提取图像的兴趣区域 (ROI)。

2.5.3 几何空间变换 (Affine Transformations)

几何变换改变图像中像素的空间排列, 通常利用齐次坐标将多种变换统一表示。坐标 (x, y) 到 (x', y') 的仿射变换由矩阵 \mathbf{A} 描述:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

下表总结了常用的几何变换、参数意义及其对应的正反向矩阵:

变换类型	参数与视觉效果说明	正向矩阵 A	反向矩阵 A^{-1}
平移	t_x : 沿 x 轴移动的距离 t_y : 沿 y 轴移动的距离	$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & -t_x \\ 0 & 1 & -t_y \\ 0 & 0 & 1 \end{bmatrix}$
旋转	θ : 绕坐标原点逆时针旋转的角度	$\begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$
缩放	c_x : 沿 x 轴伸展或压缩比例 c_y : 沿 y 轴伸展或压缩比例	$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1/c_x & 0 & 0 \\ 0 & 1/c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$
垂直偏移	s_v : 垂直方向的倾斜因子 (x 坐标受 y 的线性影响)	$\begin{bmatrix} 1 & s_v & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -s_v & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
水平偏移	s_h : 水平方向的倾斜因子 (y 坐标受 x 的线性影响)	$\begin{bmatrix} 1 & 0 & 0 \\ s_h & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 \\ -s_h & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

表 1: *
表 1: 常用几何变换及其正反向仿射变换矩阵汇总

2.5.4 灰度内插

几何变换后，映射后的坐标通常不落在原始离散像素网格的整数点上，因此必须使用内插法为新位置赋灰度值：

1. **最近邻插值 (Nearest Interpolation)**: 将离目标位置最近的像素灰度赋给该点。其计算速度最快，但在放大图像时易产生明显的锯齿状人工痕迹。
2. **双线性插值 (Bilinear Interpolation)**: 利用目标点周围 2×2 邻域内的四个像素进行双线性加权。该方法比最近邻更平滑，且计算量适中。
3. **双三次插值 (Bicubic Interpolation)**: 利用目标点周围 4×4 邻域内的 16 个像素进行运算。该方法能最好地保留图像细节和边缘，是当前商业图像软件（如 Photoshop）中的标准插值法。

2.5.5 例题 7：通用仿射变换矩阵的逆变换

题目：已知坐标的仿射变换由如下齐次坐标矩阵形式给出：

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix}$$

其坐标映射关系为：

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

为了实现逆映射，需要求出反变换矩阵 \mathbf{A}^{-1} ，使得 $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{A}^{-1} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$ 。请给出 \mathbf{A}^{-1} 的详细解析表达式。

解答：

1. 建立线性方程组：根据矩阵乘法，变换关系可以展开为如下两个线性方程：

$$\begin{cases} x' = a_{11}x + a_{12}y + a_{13} \\ y' = a_{21}x + a_{22}y + a_{23} \end{cases}$$

我们的目标是解出 x 和 y 关于 x' 和 y' 的函数。

2. 计算行列式：设左上角 2×2 矩阵的行列式为 Δ （它决定了变换是否可逆，通常代表缩放和旋转的组合效果）：

$$\Delta = a_{11}a_{22} - a_{12}a_{21}$$

3. 利用矩阵求逆公式：对于形如 $\mathbf{A} = \begin{bmatrix} \mathbf{M} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ 的分块矩阵，其逆矩阵为 $\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{M}^{-1} & -\mathbf{M}^{-1}\mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ 。

通过代数运算（或克莱姆法则）解方程组，可得反变换矩阵：

$$\mathbf{A}^{-1} = \frac{1}{\Delta} \begin{bmatrix} a_{22} & -a_{12} & a_{12}a_{23} - a_{22}a_{13} \\ -a_{21} & a_{11} & a_{21}a_{13} - a_{11}a_{23} \\ 0 & 0 & \Delta \end{bmatrix}$$

4. 最终反变换表达式：将系数 $1/\Delta$ 乘入矩阵，得到最终的坐标反变换矩阵：

$$\mathbf{A}^{-1} = \begin{bmatrix} \frac{a_{22}}{\Delta} & -\frac{a_{12}}{\Delta} & \frac{a_{12}a_{23} - a_{22}a_{13}}{\Delta} \\ -\frac{a_{21}}{\Delta} & \frac{a_{11}}{\Delta} & \frac{a_{21}a_{13} - a_{11}a_{23}}{\Delta} \\ 0 & 0 & 1 \end{bmatrix}$$

参数物理意义总结：

- a_{13}, a_{23} ：在正向变换中负责平移。在反变换矩阵中，它们被重新组合以抵消平移的影响。
- a_{11}, a_{22} ：主要负责尺度伸缩。在反变换中，它们的位置互换并除以行列式，起到“反向伸缩”的作用。
- a_{12}, a_{21} ：主要负责旋转和剪切。在反变换中，它们的符号变反，用于撤销角度偏转。

2.6 第二章核心重点问题总结

1. **为什么要研究人类视觉系统？**因为数字图像的质量评价多由人眼主观判断，且许多图像技术（如屏幕刷新率、量化级别、图像压缩等）需根据人眼生理特性设计，了解人类认知机制对开发新算法至关重要。
2. **数字图像是如何形成的？其数学模型是什么？**图像由感知器官接收物体的反射或透射能量形成。数学上，图像可用二维函数 $f(x, y)$ 表示，其值由入射分量 $i(x, y)$ （照射源能量）和反射分量 $r(x, y)$ （或透射系数）的乘积决定，即 $f(x, y) = i(x, y)r(x, y)$ 。
3. **什么是图像的取样与量化？**取样是将连续图像的坐标进行数字化的过程，决定了图像的空间分辨率。量化是将图像函数幅值（灰度值）进行数字化的过程，决定了图像的灰度分辨率。二者共同将连续图像转换为由像素组成的数字阵列。
4. **什么是空间分辨率和灰度分辨率？**空间分辨率是图像中可辨别的最小细节的度量，由取样点数（如 $M \times N$ ）决定。灰度分辨率指在灰度级中可分辨的最小变化，通常取为 2 的整数次幂（如 $L = 2^k$ ），级数越多，图像质量越好但存储量越大。
5. **像素间存在哪些基本关系？**包括邻域关系（4-邻域 N_4 、D-邻域 N_D 、8-邻域 N_8 ）、邻接与连通性（4-连通、8-连通、m-连通，用于区分边界和区域）以及距离度量（欧氏距离 D_e 、城市距离 D_4 、棋盘距离 D_8 ）。
6. **图像处理中常用的几何变换和内插方法有哪些？**常用几何变换为仿射变换，包括平移、旋转、缩放、偏移等。执行变换后需通过灰度内插为新像素赋值，常见方法有计算最快的最近邻插值、平滑度较好的双线性插值以及保留细节最佳的双三次插值。
7. **图像算术运算有哪些典型应用？**主要包括利用图像均值（加法）进行图像去噪，利用图像相减（减法）进行差异增强（如 DSA 血管造影），以及利用乘除法进行阴影校正或 ROI 模板提取。