

FDU 数字图像处理 1. An Introduction

本文参考以下教材:

- Digital Image Processing (4th Edition, R. Gonzalez, R. Woods) Chapter 1, 2
- 数字图像处理 (第四版, R. Gonzalez, R. Woods) 第 1, 2 章

欢迎批评指正!

1.1 什么是数字图像处理

一幅**图像** (image) 可以定义为一个二维函数 $f(x, y)$, 其中 x 和 y 是平面坐标.

任意一对平面坐标 (x, y) 处的幅值 $f(x, y)$ 称为图像在该点的**强度** (intensity) 或**灰度** (gray level)

当 (x, y) 和灰度值 $f(x, y)$ 都是有限的离散量时, 我们称该图像为**数字图像** (digital image)

其每个元素都有特定的位置和数值, 称为图像元素 (image elements), 简称**像素** (pixel)

数字图像处理 (digital image processing) 就是指借助计算机对数字图像的处理:

- 一类是其输入和输出都是图像的处理方法, 例如图像增强和图像复原
- 另一类是其输入是图像, 但输出是从输入图像中提取的属性的方法, 例如图像的特征提取和目标识别.

1.2 数学基础

1.2.1 邻接

坐标 (x, y) 处的像素 p 具有 2 个水平的相邻像素 (坐标为 $(x-1, y)$ 和 $(x+1, y)$)

以及 2 个垂直的相邻像素 (坐标为 $(x, y-1)$ 和 $(x, y+1)$)

这组像素称为像素 p 的 4-邻域, 记为 $N_4(p)$

- 若像素 $q \in N_4(p)$, 则称像素 p, q 是 **4-邻接的** (4-adjacent).

像素 p 的 4 个对角相邻元素 (坐标为 $(x-1, y-1), (x-1, y+1), (x+1, y-1), (x+1, y+1)$) 记为 $N_D(p)$

$N_4(p)$ 和 $N_D(p)$ 合并就构成了像素 p 的 8-邻域, 记为 $N_8(p)$

- 若像素 $q \in N_8(p)$, 则称像素 p, q 是 **8-邻接的** (8-adjacent).

为消除 8-邻接概念可能导致的歧义, 我们引入**混合邻接** (mixed adjacency) 的概念:

- 对于某个给定的像素集合 V 中的像素 p, q
若 $q \in N_4(p)$ 或 $\begin{cases} q \in N_D(p) \\ (N_4(p) \cap N_4(q)) \cap V = \emptyset \end{cases}$, 则称 p, q 是混合邻接的.

以下图为例, 图 (b) 最上面的三个像素显示了多个 8-邻接

若限定为混合邻接, 则变为图 (c) 所示的情况:

(a)像素的排列; (b)8 邻接像素 (邻接由虚线显示); (c)m 邻接;

像素序列 $(x_0, y_0), \dots, (x_n, y_n)$ 构成**通路** (path)

当且仅当对于任意 $1 \leq i \leq n$ 像素 (x_{i-1}, y_{i-1}) 和 (x_i, y_i) 都是邻接的, 此时称 n 为**通路长度** (length).

若通路的首尾像素相同 (即 $(x_n, y_n) = (x_0, y_0)$), 则称它是**闭合的** (closed)

我们可以根据指定的邻接类型定义 4-邻接通路、8-邻接通路和混合邻接通路.

(注意在下列定义中, 我们使用的是 4-邻接或 8-邻接)

设 S 表示图像中像素的一个子集.

若像素 $p, q \in S$ 之间存在一个完全由 S 中像素构成的通路, 则称 p, q 在 S 中是**连通的** (connected)

对于任意 $p \in S$, 我们称 S 中连通到该像素的像素集为 S 的一个包含 p 的**连通分量** (connected component)

若 S 仅有一个连通分量 (即其所有像素之间都是连通的), 则称 S 为**连通集** (connected set), 又称图像的一个**区域** (region)

若图像中的区域 R_1, R_2 的并集是一个连通集, 则称 R_1, R_2 是**邻接区域** (adjacent region)

不邻接的区域称为**不相交区域** (disjoint region)

1.2.2 距离

我们称 $d(\cdot, \cdot)$ 是一个**距离** (distance), 如果对于任意像素 p, q, s 它都满足:

- ① 非负性和正定性: $d(p, q) \geq 0$ 当且仅当 $p = q$ 时取等
- ② 对称性: $d(p, q) = d(q, p)$
- ③ 次可加性: $d(p, s) \leq d(p, q) + d(q, s)$

我们常用的距离有:

(设 p, q 的坐标为 $(x, y), (u, v)$)

- Euclid 距离: $d_2(p, q) = \sqrt{(x - u)^2 + (y - v)^2}$
- Manhattan 距离: $d_1(p, q) = |x - u| + |y - v|$
- 棋盘距离: $d_\infty(p, q) = \max\{|x - u|, |y - v|\}$

1.2.3 算子

考虑一般的算子 \mathcal{H} , 它根据输入图像 f 产生一幅图像 $\mathcal{H}(f)$

若对于任意两幅图像 f_1, f_2 和常数 a, b 都有 $\mathcal{H}(af_1 + bf_2) = a\mathcal{H}(f_1) + b\mathcal{H}(f_2)$ 成立,

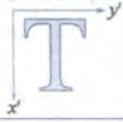

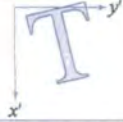
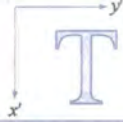
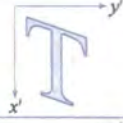
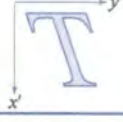
则我们称 \mathcal{H} 是一个**线性算子** (linear operator), 否则称为**非线性算子**

线性运算非常重要, 因为它们包含了大量适用于图像处理的理论与实践成果.

非线性运算的范围非常有限, 但在某些场景中它们的性能远优于线性运算.

1.2.4 空间运算

表 2.3 基于式(2.45)的仿射变换

变换名称	仿射矩阵 A	坐标公式	示 例
恒等	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x$ $y' = y$	
缩入/反射 (对于反射, 将一个比例因子设为-1, 而将另一个比例因子设为0)	$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = c_x x$ $y' = c_y y$	
(关于原点) 旋转	$\begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x \cos \theta - y \sin \theta$ $y' = x \sin \theta + y \cos \theta$	
平移	$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x + t_x$ $y' = y + t_y$	
(垂直) 剪切	$\begin{bmatrix} 1 & s_v & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x + s_v y$ $y' = y$	
(水平) 剪切	$\begin{bmatrix} 1 & 0 & 0 \\ s_h & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x$ $y' = s_h x + y$	

1.2.5 逻辑运算

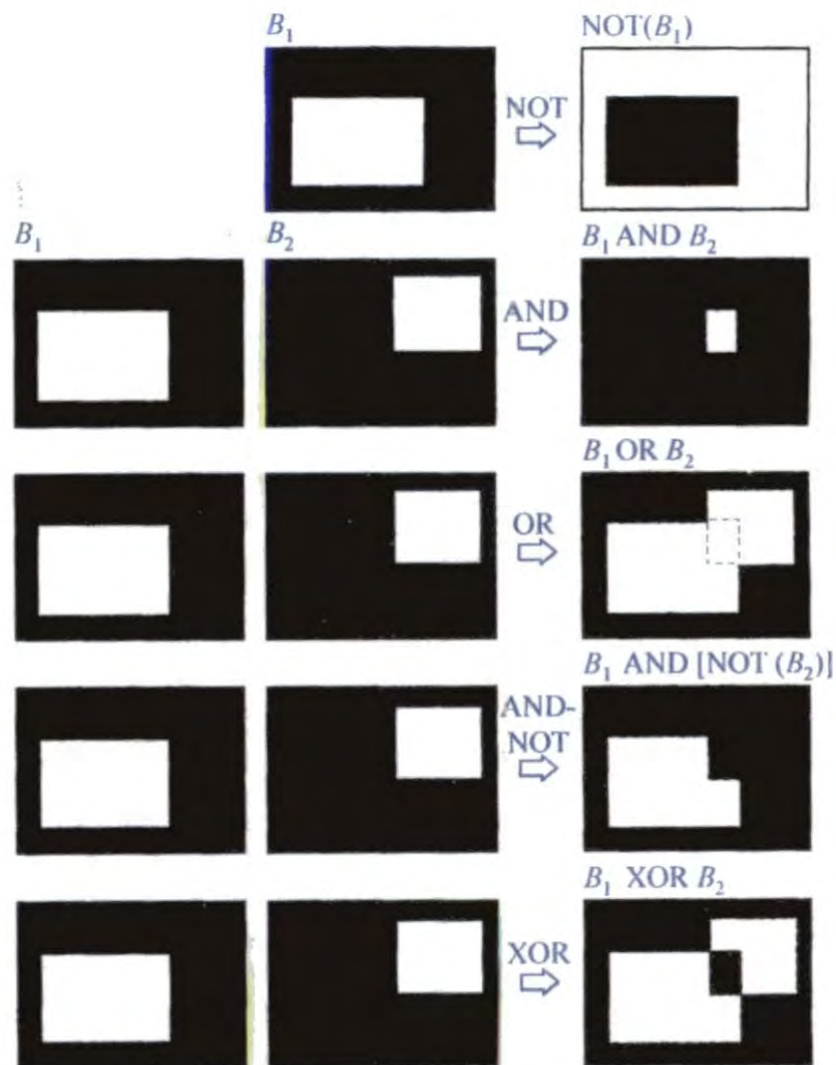


图 2.37 涉及前景（白色）像素的逻辑运算的说明。黑色表示 0，白色表示 1。虚线仅作为参照，不是结果的一部分

1.3 基本概念

对比度 (contrast):

$$\text{Contrast} := \frac{\text{Luminance difference}}{\text{Average luminance}}$$

Weber 对比度主要用于评价目标 (光强为 I) 在背景 (光强为 I_b) 上是否容易被检测到:

$$C_{\text{weber}} := \frac{I - I_b}{I_b}$$

对于一个亮的目标在暗的背景上，Weber 对比度可以用来衡量目标的显著性.

Michelson 对比度常用于周期性波形或图像的对比度指标，它描述了最大亮度和最小亮度之间的相对差异:

$$C_{\text{Michelson}} := \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}$$

