

数字图像基础

本章从连续图像的数学描述入手,通过分析连续图像与数字图像的不同,介绍从连续的模拟图像转换为离散的数字图像的取样与量化原理和方法。阐述了常用的数字图像文件格式与数据结构,为处理数字图像提供基本的知识,同时对数字成像设备原理和灰度直方图的概念、性质应用进行了阐述。

3.1 连续图像的数学描述

在第1章和第2章已经分别讨论过图像的模型,在此,把它们放在一起进一步讨论。

一幅三维空间的图像可以被看作是空间上各点光强度的集合,它随图像坐标 (x, y, z) 、光线的波长 λ 和时间 t 的变化而变化。其数学表达式为

$$g = f(x, y, z, \lambda, t) \quad (3.1.1)$$

对于二维平面图像而言,其表示为

$$g = f(x, y, \lambda, t) \quad (3.1.2)$$

若只考虑光的能量而不考虑其波长,图像在视觉上则表现为灰色影像,称之为灰度图像或单色图像,其图像函数为

$$g = f(x, y, t) = \int_0^{\infty} f(x, y, \lambda, t) V_s(\lambda) d\lambda \quad (3.1.3)$$

式中, $V_s(\lambda)$ 为相对视敏函数。

若考虑不同波长光的彩色效应,则图像在视觉上表现为彩色图像,其图像函数为

$$g = \{f_R(x, y, t), f_G(x, y, t), f_B(x, y, t)\} \quad (3.1.4)$$

式中

$$f_R(x, y, t) = \int_0^{\infty} f(x, y, \lambda, t) R_s(\lambda) d\lambda \quad (3.1.5)$$

$$f_G(x, y, t) = \int_0^{\infty} f(x, y, \lambda, t) G_s(\lambda) d\lambda \quad (3.1.6)$$

$$f_B(x, y, t) = \int_0^{\infty} f(x, y, \lambda, t) B_s(\lambda) d\lambda \quad (3.1.7)$$

式中, $R_s(\lambda)$ 、 $G_s(\lambda)$ 、 $B_s(\lambda)$ 依次为红、绿、蓝三原色的视敏函数。

图像内容随时间变化的图像称为运动图像,而图像内容不随时间变化的图像称为静止

图像。本书图像处理部分主要研究静止图像,图像通信部分会涉及到运动图像。静止图像一般用 $f(x,y)$ 表示。因为光是能量的一种形式,故 $f(x,y) > 0$,即

$$0 < f(x,y) < \infty \quad (3.1.8)$$

在每天的视觉活动中,人眼看到的图像一般都是由物体反射的光组成的。 $f(x,y)$ 可被看成由两个分量组成:一个分量是所见场景的入射光量,另一个分量是场景中物体反射光量的能力,即反射率。这两个分量被称为照射分量和反射分量,分别用 $i(x,y)$ 和 $r(x,y)$ 表示。这就意味着对于反射光形成的图像 $f(x,y)$,满足

$$f(x,y) = i(x,y)r(x,y) \quad (3.1.9)$$

第2章已给出了一些典型值,请参照第2章相关内容。

3.2 图像的取样与量化

第2章曾介绍过几种获取图像的方法,所有这些方法的目的是相同的,就是通过传感器从感知的数据中产生图像。这些传感器和人眼的视网膜成像基本原理一致,它们获得的图像信号一般是连续的,称之为模拟图像。这种图像的连续性包含两个方面的含义,即空间位置延续的连续性和每一个位置上光的强度变化的连续性,因而产生的图像为连续的模拟函数。然而,用连续的模拟函数表示的图像无法使用计算机进行处理,也无法直接在各种数字系统中传输或存储,必须将代表图像的连续(模拟)信号转变为空间位置和强度变化均不连续的离散(数字)信号。对于可见光而言,光强是连续的,因此大多数传感器输出的是连续电压波形,不符合数字图像处理的要求。因此,为了产生一幅数字图像,需要把连续的感知数据转换为离散的数字形式。这必须包括两个步骤进行处理,即取样和量化。

图3.1是连续图像及其对应的离散图像示意图。

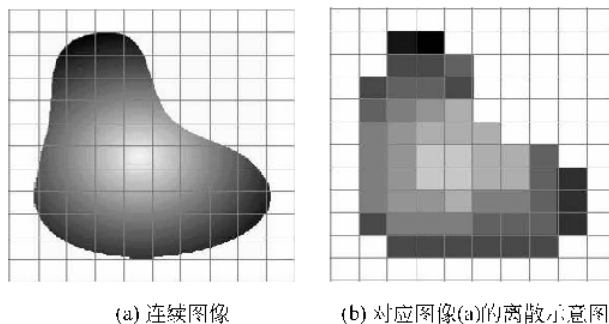


图 3.1 连续图像与离散图像

3.2.1 取样

取样是指将空间域上或时域上连续的图像(模拟图像)变换成离散采样点(像素)集合的一种操作,也叫做抽样、采样。被选取的点称为取样点、抽样点、采样点或样点,在二维数字图像中,这些样点被称为像素(pixel)。在取样点上的函数值称为取样值、抽样值、采样值或

样值。通过取样操作就能在空间上用有限的取样点来代替连续无限的坐标值。

一幅图像应取多少样点需要综合考虑,取样点过多,会增加用于表示这些取样点的信息量,传输和存储的要求就更高;如果取样点过少,则很可能丢失原图像所包含的信息,使数字化后的图像失去细节,画面粗糙。所以,应该找到一个合适的取样值,即采用某种方法能够完全重建图像,又不至于增加额外的存储和传输需求,这需要用二维取样定理来解决。

由于图像基本上是采取二维平面信息的分布方式来描述的,所以为了对它进行取样操作,需要先将二维信号变为一维信号,再对一维信号完成取样。换句话说,将二维取样转换成两次一维取样操作来实现。

图 3.2 是图像取样处理过程示意图。将二维图像信号变换成一维图像信号最常用的方法是,首先沿垂直方向按一定空间间隔,从上到下的顺序沿水平方向以直线扫描的方式,取出各个水平

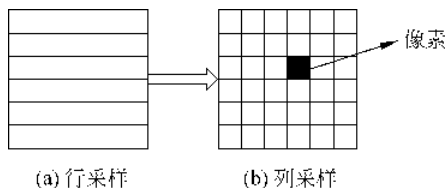


图 3.2 二维图像取样示意图

再对一维扫描线信号按一定间隔取样得到离散信号。换句话说,图像取样是通过先在垂直方向上取样,然后将得到的结果再沿水平方向取样两个步骤来完成的。经过取样之后得到的二维离散信号可用一个矩阵来表示,其最小单位即矩阵当中的每个值就代表一个取样点的值,称为像素。一般情况下,水平方向的取样间隔与垂直方向的取样间隔相同。

对于运动图像(即时域上连续的图像),首先在时间轴上取样,其次沿垂直方向取样,最后沿水平方向取样。

对一幅图像取样后,若每行(即横向)像素为 M 个,每列(即纵向)像素为 N 个,则图像大小为 $M \times N$ 个像素。例如,一幅 640×480 的图像,就表示这幅连续图像在长、宽方向上分别分成 640 个像素和 480 个像素。显然,想要得到更加清晰的图像质量,就要提高图像的取样像素点数,也就是要使用更多的像素点来表示该图像,但相对需要付出更大的存储空间。

3.2.2 二维取样定理

通过取样可以将连续的图像信号数字化,但在取样时,我们只选用了非常有限的一些离散的点,这些点并不一定能准确地表示原有的图像信号,那么由离散点形成的数字图像信号怎样才能精确地恢复原图像所需要的最少信息呢?如果取样间隔过大,图像的分辨率必然较低,不能精确恢复原图像;如果取样间隔太小,虽然能够很好地恢复原图像,但存储空间必然很大。因此,从另一个角度理解就是要确定各取样点在水平和垂直方向的最大间隔,从而在确保精确恢复原图像的同时使用最少的存储空间,这一问题由二维取样定理解决。

由于二维取样定理涉及傅里叶频谱,因此首先介绍一下二维信号的傅里叶变换。我们知道,对于一维有界信号 $f(t)$,其傅里叶变换和傅里叶反变换分别定义为

$$F(f) = \int_{-\infty}^{\infty} f(t) e^{-j2\pi ft} dt \quad (3.2.1)$$

$$f(t) = \int_{-\infty}^{\infty} F(f) e^{j2\pi ft} df \quad (3.2.2)$$

其中, $F(f)$ 被称为 $f(t)$ 的频谱。其物理意义是, $f(t)$ 可由时间域或空间域上的各谐波分量叠加得到。更详细的物理意义在第 4 章阐述。

在二维情况下, 类似的定义 $f(x, y)$ 的傅里叶变换 $F(u, v)$, 其傅里叶变换与傅里叶反变换分别定义为

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy \quad (3.2.3)$$

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{j2\pi(ux+vy)} du dv \quad (3.2.4)$$

$F(u, v)$ 被称为 $f(x, y)$ 的频谱, 同样, 它的物理意义是空间域上的各谐波分量的叠加, 表明频率成分与二维图像信号之间的相互关系。

对于待处理的实际二维图像, 其傅里叶变换一般都是在频率域上有界的, 也就是傅里叶变换后的频率总落在一定的频率域范围之内。

奈奎斯特(Nyquist)指出, 如果时域信号 $f(t)$ 的频谱 $F(f)$ 的上界为 f_m , 那么只要取样点的间隔 $\Delta t \leq 1/(2f_m)$, 时域信号就可被精确地恢复, 这就是一维奈奎斯特取样定理。将一维奈奎斯特取样定理推广, 即为二维取样定理, 其具体描述如下。

对于二维有限带宽的模拟图像信号 $f(x, y)$, 如果其二维傅里叶谱 $F(u, v)$ 在水平方向的截止频率为 U_m , 在垂直方向的截止频率为 V_m , 即傅里叶变换只在 $|v| \leq V_m$ 和 $|u| \leq U_m$ 的范围内不为 0, 那么只要水平方向的空间取样频率 $U_0 \geq 2U_m$, 垂直方向的空间取样频率 $V_0 \geq 2V_m$, 即水平取样间隔满足 $\Delta x \leq 1/2U_m$, 垂直取样间隔 $\Delta y \leq 1/2V_m$, 图像就可被精确地恢复。

在理解二维取样定理时, 可以用一个例子说明。例如, 一个镜头可以在焦平面成像为 50 线对/毫米的正弦光栅, 只要图像传感器的 CCD 或 CMOS 的排列密度达到每毫米 100 个单元, 则形成的数字图像就可以完全精确地记录该影像信息, 即达到镜头的最高分辨率。这时再提高传感器的排列密度实际上是没有更大实际意义的。如果传感器的排列密度不能达到上述要求, 则会因为频率混叠效应, 造成画面上细节信息的模糊。

3.2.3 取样对数字图像效果评价的影响

连续图像经过取样之后所获得的数字图像的效果与以下几个评价参数有关。

- 图像分辨率: 取样所获得的图像总像素的多少, 通常称为图像分辨率, 也叫空间分辨率。例如, 640×480 图像的总像素数为 307 200 个, 在购买具有这种分辨率的数字化设备时, 产品性能介绍上会给出 30 万像素分辨率这一参数。
- 采样密度: 是指在图像上单位长度所包含的取样点数。取样密度的倒数是像素间距。
- 取样频率: 是指 1s 内取样的次数。它反映了取样点之间的间隔大小, 取样频率越高, 丢失的信息越少, 取出的样本越细腻逼真, 图像的质量越好, 但要求的存储量也就越大。
- 扫描分辨率: 表示一台扫描仪输入图像的细微程度, 指每英寸(1 英寸 ≈ 2.5 厘米)扫描所得到的点, 单位是 DPI(dot per inch)。数值越大, 表示被扫描的图像转化为数字化图像越逼真, 扫描仪质量也越好。

无论采用哪种评价参数,实际上在进行取样时,取样点间隔的选取是一个非常重要的问题,它决定取样后图像的质量,即忠实于原图像的程度。取样间隔的大小取决于原图像中包含的细微明暗变化。

图 3.3 分别显示了对 Lena 的同一幅图像按 128×128 、 64×64 、 32×32 、 16×16 以及 8×8 个取样点进行取样的结果。从采样的结果很容易分辨出质量好坏与采样像素点多少的关系。

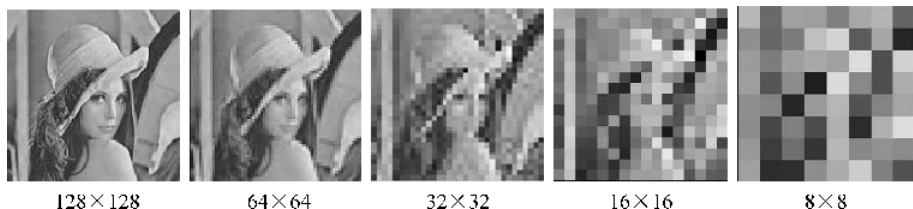


图 3.3 以不同分辨率取样的结果

3.2.4 图像的量化和编码

这里所述的图像是指静态的图像。对动态图像的量化,实际上也是建立在对多个静态图像量化的基础上的。

图像的数字化不仅包括空间上的离散化(取样),还包括取样点函数值的离散化。经取样后图像被分割成空间上离散的像素,但其灰度是连续的,还是一个有无穷多个取值的连续变化量,不能直接用于计算机进行处理。因此,必须将其转化为有限个离散值并赋予不同的码字才能真正成为数字图像,这个将像素连续灰度值转换成离散的整数值的過程叫量化。

图 3.4 说明了取样和量化的基本概念。图 3.4(a)显示了一幅连续图像 $f(x, y)$, 可以把它转换为数字形式。一幅图像的 x 和 y 坐标及幅度可能都是连续的,为了把它转换为数字形式,必须在坐标位置和幅度上都做取样操作。数字化坐标值称为取样,数字化幅度值称为量化。

图 3.4(b)显示的一维函数是图 3.4(a)中沿线段 AB 的连续图像幅度值(灰度级)的曲线。随机变化是由图像噪声引起的。沿 AB 线段等间隔对该函数取样,如图 3.4(c)所示。每个取样位置由图底部的垂直小线段给出,取样点的位置如函数曲线上的白色小方块所示,这一组离散位置给出了取样函数。然而,取样值在灰度级值上还是连续分布的。为了形成数字函数,灰度值也必须转变为离散值(数字化)。图 3.4(c)的右侧显示了分为 8 个离散级别的灰度级标尺,其范围从黑到白,垂直面上的小线段标记指示赋予 8 级灰度的每一级一个特定值。将每一取样的连续灰度值简单量化为 8 个离散灰度级中的一个,处于两个等级之间的灰度赋值时,必须将样值对标记在垂直方向上的等级做近似,更接近于哪一级就赋予哪一级的离散灰度等级。由取样和量化产生的数字样值如图 3.4(d)所示,从图像的顶点开始逐行执行这一过程就能得到一幅二维数字图像。

经量化后的数字图像中不同灰度值的个数称为灰度级数,用 G 表示,图 3.4 的灰度级数只有 8 个。若一幅数字图像的量化灰度级数 $G=256=2^8$ 级,则像素灰度取值范围是 $0 \sim 255$ 之间的整数。由于用 8b 就能表示灰度图像像素的灰度值,因此常称为 8b 量化。数字

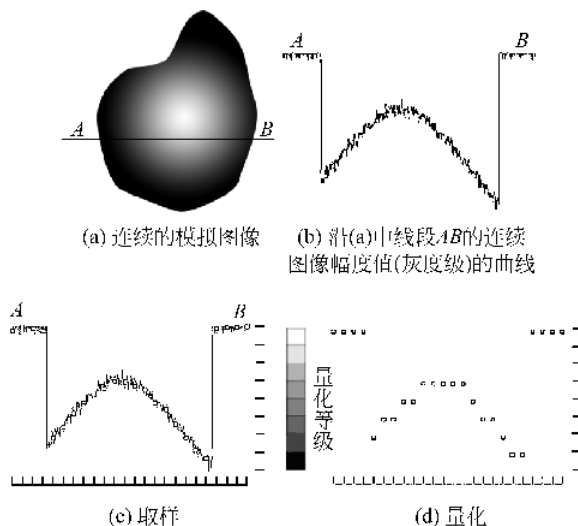
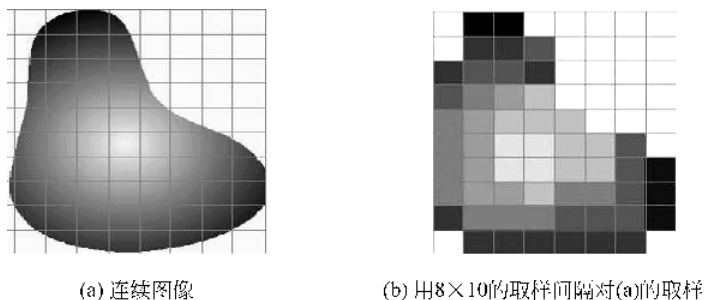


图 3.4 取样和量化的基本概念

化前需要决定图像大小(行数 M 、列数 N)和灰度级数 G 的取值。图 3.5 是对一个连续图像的取样和量化示意图,其中图 3.5(a)是一幅连续图像 $f(x, y)$,图 3.5(b)是用 8×10 的取样间隔对连续图像的取样,图 3.5(c)是用 256 级灰度对离散后的图 3.5(b)所示图像中各像素点进行量化的结果。



G

255	0	0	255	255	255	255	255
255	46	11	56	255	255	255	255
11	58	56	32	255	255	255	255
62	117	160	203	255	255	255	255
124	162	219	232	229	255	255	255
138	168	243	248	231	218	100	255
138	166	255	250	226	213	73	39
130	160	180	215	144	129	63	13
11	111	122	121	68	59	55	0
255	7	33	36	32	30	44	220

(c) 用 256 级灰度对 (b) 中各像素的量化结果

图 3.5 取样与量化示意图

由于一般数字图像灰度级数 G 为 2 的整数幂,即 $G=2^n$,那么一幅大小为 $M \times N$ 、灰度级数为 G 的图像所需的存储空间 $M \times N \times n$ (比特)称为图像的数据量。由此可以知道,当图像量化的等级越多,需要的数据量就越大。怎样确定数字图像的灰度等级,这就需要了解量化所带来的误差和失真。

由于量化是以有限个离散值来近似表示无限多个连续量,因此一定会产生误差,这就是所谓的量化误差。由量化误差所产生的失真叫做量化失真,或称引入了量化噪声。当量化层次较多时,量化值与实际的模拟值之间的差值较小,量化误差不大,引起的量化失真比较轻微,以至于人眼不能分辨。而当量化层次少到一定程度时,量化值与模拟值之间的差值(量化误差)变得很严重,引起严重的图像量化失真,尤其在原来亮度值缓慢变化的区域将引起生硬的阶梯状伪轮廓。图像量化的基本要求就是在量化噪声对图像质量的影响可忽略的前提下用最少的量化层次进行量化。

从视觉效果来看,采用大于或等于 6 比特(64 级)量化的灰度图像,视觉效果就能令人满意。因为计算机中 1 字节刚好是 8 比特,因此对于图像灰度等级的量化,通常采用 8 比特、4 比特、2 比特或 1 比特等方式。图 3.6 是分别用 8 比特(256 级)、4 比特(16 级)、2 比特(4 级)和 1 比特(2 级)对分辨率为 128×128 的 Lena 图像量化的结果。由图可知,量化所采用的位数越少,引起的量化失真越严重。特别是只用 1 比特或 2 比特进行量化时,大量的图像细节已经被噪声掩盖,图像已经严重失真。



图 3.6 分辨率为 128×128 、不同量化等级的 Lena 图像量化结果

由于量化层数 G 为 2 的 n 次幂,因此可对应地用 n 比特自然二进制码表示,形成最通用的 PCM 编码。均匀量化是数字化时通常采用的量化方式,如果采用等间隔分层,那么量化的层次越多,量化误差就越小,但是编码时占用的位数就会越多。例如采用 8 比特量化,则图像灰度等级就分为 256 级,即最白处用 255 表示,最黑处用 0 表示。例如,某像素点的灰度模拟值为 192.4 时,则将其量化为 192,可用二进制码 11000000 来表示。

在对像素点进行 n 比特的线性 PCM 编码时,每个量化分层间隔(量化步长)的相对值为 $1/2^n$ 。假定取样值在它的动态范围内的概率分布是均匀的,则可以证明,量化误差的均方值为式(3.14)中的 N_q

$$N_q = \left(\frac{1}{12}\right) \left(\frac{1}{2^n}\right)^2 \quad (3.2.5)$$

由此可得,信号峰值功率 S_{pp} 与量化误差的均方值 N_q 之比为

$$\frac{S_{pp}}{N_q} = 10 \lg 12 \cdot (2^n)^2 \approx 10.8 + 6n \text{ dB} \quad (3.2.6)$$

上式表征了线性 PCM 编码的性能,通常将其简称为量化信噪比,并用 $(S/N)_q$ 表示。

由式(3.2.6)可知,每个像素的编码比特数 n 直接关系到数字化的图像质量,每增减 1 比特,就会使量化信噪比增减约 6dB。选择量化的比特数 n 经常用主观评价的方法,比较原图像与量化图像的差别,或者直接观看量化后的图像,当量化引起的误差已不能被人眼所觉察或在人眼可接受的范围内时,所对应的最小量化比特数即为 n 。考虑到采用大于或等于 6 比特(64 级)量化的灰度图像,视觉效果就能基本令人满意,并且一个字节为 8 比特,因此对于一般的应用,如普通质量的静态图像、广播电视和视频通信等,均采用 8 比特进行量化和编码。对于一些图像质量要求较高,特别是强调亮部和暗部细节的图像,如高质量的静止图像、遥感图像处理等,需要 10 比特或更高精度的图像。目前使用的数码相机中最高的量化层次已采用 14 比特进行量化。

在上述的量化过程中,实际上是对每个像素点的灰度值进行等间隔的层次划分,这叫做均匀量化。有时为了突出图像的亮部细节或暗部细节,还会采用另一种非等间隔的层次划分的非均匀量化。

3.2.5 数字图像表示

取样和量化的结果实际是一个矩阵。在本书中将用两种主要方法来表示数字图像。假如一幅图像 $f(x, y)$ 被取样,则产生的数字图像有 M 行和 N 列。经量化后,坐标 (x, y) 的值已经变成离散量。为表达清楚和方便起见,对这些离散的坐标采用整数表示。这样,原点的坐标值是 $(x, y) = (0, 0)$,沿图像第一行的下一个坐标值用 $(x, y) = (0, 1)$ 来表示。用符号 $(0, 1)$ 表示沿第一行的第二个取样是很重要的,但它并不意味着是图像取样的物理坐标实际值。图 3.7 显示了贯穿本书所用的坐标约定。

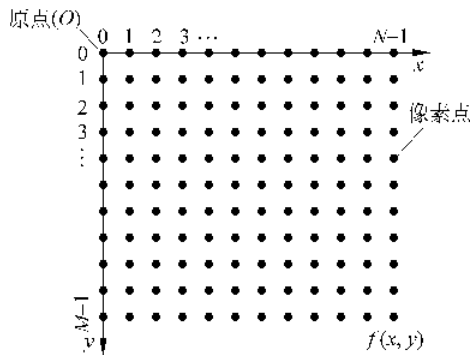


图 3.7 图像坐标

上面介绍的表示法允许以下面的紧凑矩阵形式写出完整的 $M \times N$ 数字图像。

$$f(i, j) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix} \quad (3.2.7)$$

这个表达式的右侧定义了一幅数字图像。矩阵中的每个元素称为图像单元、图像元素或像素,图像和像素这两个术语将在我们讨论的其余部分用以表示数字图像和它的元素。

在某些讨论中,可以用传统矩阵表示法来表示数字图像和像素。

$$\mathbf{A} = \begin{bmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,N-1} \\ \vdots & \vdots & & \vdots \\ a_{M-1,0} & a_{M-1,1} & \cdots & a_{M-1,N-1} \end{bmatrix} \quad (3.2.8)$$

显然, $a_{ij} = f(x=i, y=j) = f(i, j)$, 因此, 式(3.2.7)和式(3.2.8)是相同矩阵的两个不同表达形式。

以更正规的数学术语描述取样和量化有时可能有用。令 $\mathbf{Z}$ 和 $\mathbf{R}$ 分别表示整数集和实数集。取样过程可看做是把 xy 平面分为网格, 每一网格中心的坐标是笛卡儿积 Z^2 的一对元素, Z^2 是所有有序元素对 (z_i, z_j) 的集合, z_i 和 z_j 是 Z 中的整数。因此, 如果 (x, y) 是 Z^2 中的整数, 并且 f 是把灰度级值(即实数集 $\mathbf{R}$ 的一个实数)赋予每个特定坐标对 (x, y) 的函数, 则 $f(x, y)$ 就表示一幅数字图像。很显然, 这个函数赋值过程就是前边描述的量化过程。如果灰度也是整数, 则用 $\mathbf{Z}$ 代替 $\mathbf{R}$, 数字图像变成一个二维函数且它的坐标和幅值都是整数。

3.2.6 像素间的一些基本关系

为更好地表示图像中像素点间的相互关系, 以便于后续章节恰当地阐述像素间的相对位置, 对像素间的一些重要关系进行适当描述。如前边提到的那样, 一幅图像用 $f(x, y)$ 表示。当指特殊像素时, 用小写字母(如 p 和 q)表示。

1. 相邻像素

位于坐标 (x, y) 处的一个像素 p 有 4 个水平和垂直的相邻像素, 其坐标由下式给出

$$(x-1, y), (x+1, y), (x, y-1), (x, y+1)$$

这 4 个像素依次邻接于像素 p 的左、右、上和下。该像素集称为 p 的 4 邻域, 用 $N_4(p)$ 表示。每个像素距 (x, y) 一个单位距离, 如果 (x, y) 位于图像的边界, 则 p 的某一邻近像素位于数字图像的外部, 实际不存在。

p 的 4 个对角邻像素有如下坐标

$$(x-1, y-1), (x-1, y+1), (x+1, y-1), (x+1, y+1)$$

这 4 个像素依次邻接于像素 p 的左上、左下、右上和右下, 并用 $N_D(p)$ 表示。 p 的 4 对角邻域 $N_D(p)$ 与 4 邻域 $N_4(p)$ 一起组成的像素集叫做 p 的 8 邻域, 用 $N_8(p)$ 表示。与前边一样, 如果 (x, y) 位于图像的边界, 则 $N_D(p)$ 和 $N_8(p)$ 中的某些点将落入图像所能表示的边界之外。

2. 邻接性、连通性、区域和边界

像素间的连通性是一个基本概念, 它简化了许多数字图像概念的定义, 如区域和边界。为了确定两个像素是否连通, 必须确定它们是否相邻及它们的灰度值是否满足特定的相似性准则(或者说, 它们的灰度值是否相等)。例如, 在具有 0, 1 值的二值图像中, 两个像素可能是 4 邻接的, 但是, 仅仅当它们具有同一灰度值时才能说是连通的。

令 V 是用于定义邻接性的灰度值集合。在二值图像中, 如果把具有 1 值的像素归入邻接的集合, 则 $V = \{1\}$ 。在灰度图像中, 概念是一样的, 但是集合 V 一般包含更多的元素。例如, 对于具有可能的灰度值且在 0~255 范围内的像素邻接性, 集合 V 可能是这 256 个值的任何一个子集。考虑如下三种类型的邻接性。

(1) 4 邻接: 如果 q 在 $N_4(p)$ 集中, 则具有 V 中数值的两个像素 p 和 q 是 4 邻接的。

(2) 8 邻接: 如果 q 在 $N_8(p)$ 集中, 则具有 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 q 在 $N_4(p)$ 中, 或者 q 在 $N_D(p)$ 中且集合 $N_4(p) \cap N_4(q)$ 没有 V 值的像素, 则具有 V 值的像素 p 和 q 是 m 邻接的。

混合邻接是 8 邻接的改进。混合邻接的引入是为了消除采用 8 邻接常常发生的二义性。例如, 考虑图 3.8(a) 对于 $V=\{1\}$ 所示的像素安排。位于图 3.8(b) 上部的三个像素显示了多重(二义性)8 邻接, 如虚线指出的那样。这种二义性可以通过 m 邻接消除, 如图 3.8(c) 所示。如果 S_1 中的某些像素与 S_2 中的某些像素邻接, 则两个图像子集 S_1 和 S_2 是相邻接的。这里和下面的定义中邻接意味着 4、8 或者 m 邻接。

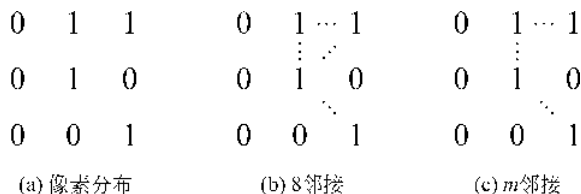


图 3.8 图像的邻接

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)是特定像素序列, 其坐标为

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

这里, $(x_0, y_0) = (x, y)$, $(x_n, y_n) = (s, t)$, 并且像素 (x_i, y_i) 和 (x_{i-1}, y_{i-1}) ($1 \leq i \leq n$) 是邻接的。在这种情况下, n 是通路的长度。如果 $(x_0, y_0) = (x_n, y_n)$, 则通路是闭合通路。可以依据特定的邻接类型定义 4、8 或 m 邻接。例如, 图 3.8(b) 所示的, 东北角和东南角点之间的通路是 8 通路, 图 3.8(c) 中的通路是 m 通路。注意, 在 m 通路中不存在二义性。

令 S 代表一幅图像中像素的子集。如果在 S 中全部像素之间存在一个通路, 则可以说两个像素 p 和 q 在 S 中是连通的。对于 S 中的任何像素 p , S 中连通到该像素的像素集叫做 S 的连通分量。如果 S 仅有一个连通分量, 则集合 S 叫做连通集。

令 R 是图像中的像素子集。如果 R 是连通集, 则称 R 为一个区域。一个区域 R 的边界(也称为边缘或轮廓)是区域中像素的集合, 该区域有一个或多个不在 R 中的邻点。如果 R 是整幅图像(设这幅图像是像素的方形集合), 则边界由图像第一行、第一列和最后一行、最后一列定义。这个附加定义是需要的, 因为这种整幅图像除了边缘没有邻点。正常情况下, 当提到一个区域时, 指的是一幅图像的子集, 并且区域边界中的任何像素(与图像边缘吻合)都作为区域边界部分全部包含于其中。

边缘的概念在涉及区域和边界的讨论中常常遇到。然而, 这些概念中有一个关键区别。一个有限区域的边界形成一条闭合通路, 边缘是由具有某些导数值(即梯度, 超过预先设定的阈值)的像素形成。这样, 边缘的概念是基于在不连续点进行灰度级测量的局部概念。把边缘点连接成边缘线段是可能的, 并且有时以与边界对应的方法连接线段, 但并不总是这样。边缘和边界吻合的一个例外就是二值图像的情况。根据连通类型和所用的边缘算子, 从二值区域提取边缘与提取区域边界是一样的。这很直观, 在概念上把边缘考虑为强度不连续的点和封闭通路的边界是有帮助的。

3. 距离度量

对于像素 p, q 和 z , 其坐标分别为 (x, y) , (s, t) 和 (v, w) 。如果:

(1) $D(p, q) \geq 0$, (当且仅当 $p = q$ 时, $D(p, q) = 0$)

(2) $D(p, q) = D(q, p)$

(3) $D(p, z) \leq D(p, q) + D(q, z)$

则 D 是距离函数或度量。

p 和 q 间的欧氏距离定义如下

$$D_e(p, q) = [(x-s)^2 + (y-t)^2]^{\frac{1}{2}} \quad (3.2.9)$$

对于距离度量, 距点 (x, y) 的距离小于或等于某一值 r 的像素是中心在 (x, y) , 半径为 r 的圆平面。

p 和 q 间的距离 D_d (也叫城市街区距离) 如下式定义

$$D_d(p, q) = |x-s| + |y-t| \quad (3.2.10)$$

在这种情况下, 距点 (x, y) 的 D_d 距离小于或等于某一值 r 的像素形成一个中心在 (x, y) 的菱形。例如, 距点 (x, y) 的 D_d 距离小于或等于 2 的像素形成固定距离的下列轮廓。

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

具有 $D_d=1$ 的像素是 (x, y) 的 4 邻域。

p 和 q 间的 D_g 距离 (也叫棋盘距离) 定义为下式

$$D_g(p, q) = \max(|x-s|, |y-t|) \quad (3.2.11)$$

在这种情况下, 距点 (x, y) 的 D_g 距离小于或等于某一值 r 的像素形成中心在 (x, y) 的方形。例如, 距点 (x, y) (中心点) 的 D_g 距离小于或等于 2 的像素形成下列固定距离的轮廓。

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

具有 $D_g=1$ 的像素是关于 (x, y) 的 8 邻域。

注意: p 和 q 之间的 D_d 和 D_g 距离与任何通路无关, 通路可能存在于各点之间, 因为这些距离仅与点的坐标有关。然而, 如果选择考虑 m 邻接, 则两点间的 D_m 距离用点间最短的通路定义。在这种情况下, 两像素间的距离将依赖于沿通路的像素值以及它们的邻点值。例如, 考虑下列安排的像素并假设 p, p_2 和 p_4 的值为 1, p_1 和 p_3 的值为 0 或 1。

$$\begin{array}{cc} p_3 & p_4 \\ p_1 & p_2 \\ p & \end{array}$$