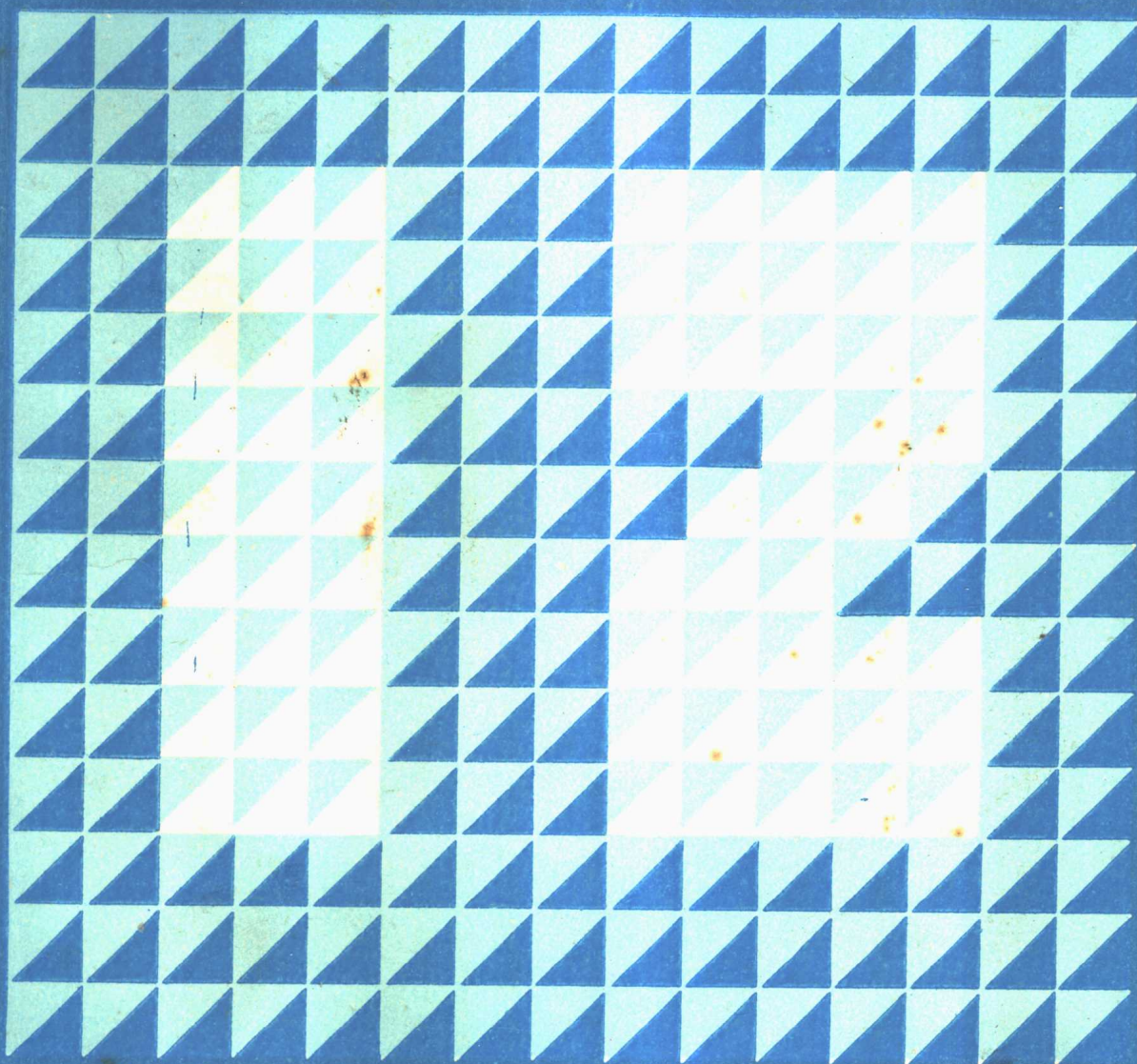


数字图象处理与模式识别

余英林 编著

华南理工大学出版社



数字图象处理与模式识别

余英林 编著

华南理工大学出版社

内 容 提 要

本书从通信理论和信号估计理论出发,系统地介绍了图象预处理、图象特征提取及图象模式识别的基本原理和各种典型应用。书中还收集了本领域国内外最新发表的重要研究成果。全书共分十三章。

本书可作为通信与电子系统、电路与信号系统及生物医学工程等有关专业的本科课程教材,也可供从事本领域工作的科研人员参阅。

数字图象处理与模式识别

余英林 编著

责任编辑 梁文厚 谢艳桂

华南理工大学出版社出版发行

(广州 五山)

各地新华书店经销

广州华南理工大学印刷厂印刷

开本 787×1092 1/16 印张 24.875 字数 574千

1990年10月第1版 1990年10月第1次印刷

印数1—1000

ISBN 7—5623—0201—4/TN·6(课)

定价: 4.90元

前 言

图象处理与模式识别是最近二十年蓬勃兴起的交叉学科,其常规操作现已成熟,所需基本设备在国际市场上也已有二至三代产品。现在,本学科已广泛应用于航空、航天、国防、公安、石油、地质、医学、工业、农业、气象、国土利用、灾害预报等方面。

十年来,我国有关单位一直在大力开展图象处理与模式识别的研究,有的通过引进设备开发新产品,有的研制基本设备,而大多数高等学校与科研单位则开展基础理论研究和应用基础研究。国家自然科学基金委员会近年也大力资助这类研究工作。目前,我国在该领域的工作已取得很大的进展。

高等学校的有关专业都开设了图象处理和模式识别的课程,其中有为硕士研究生开设的,也有为本科生开设的。因为专业课学时较少,只能讲授本学科的最基本内容,因此作者在本书中只编入最重要的基础内容,旨在使学生通过掌握这些基本概念和基本理论去做进一步的研究和开发工作。

作者在科研和教学工作中深切体会到,本学科的任何一项实用课题都涉及到图象处理与模式识别的内容,而且图象处理大多数都作为预处理,最后还必须通过图象理解和模式识别才能达到课题的要求。在国外,现有的教材和专著中大都把图象处理与模式识别分开来讲,将二者看作是二个不同的学科。在国内的一些学科划分中,则把图象处理作为信息处理的一部分而划入通信领域,而把模式识别划入自动化领域。作者以为这种划分未必妥善。事实上,图象处理与模式识别之间是有着密切的内在联系的。

作者在本书中试图从通信理论的角度把图象处理与模式识别二者联系起来,把模式分析的基本概念作为模式识别的一项重要内容收入本书,以适应交叉学科发展的趋势。

本书共分四部分,即图象预处理、图象分析、模式识别和模式分析,最后还列举了本学科的典型应用例。作者在编写过程中收集和引用了大量较新的文献资料。

本书可作为研究生和本科生高年级教材,也可供有关科技人员参考之用。

作者编写本书所依据的科研项目是由国家自然科学基金委员会资助的。

本书承蒙华南理工大学副校长徐秉铮教授推荐及暨南大学罗宇华老师审校,在此,谨致以衷心的感谢。

编 者

1988年11月于华南理工大学

目 录

第一篇 图象预处理

第一章 理论基础	(3)
§ 1-1 图象的形成及其数学表示.....	(3)
§ 1-2 图象函数与线性系统.....	(8)
§ 1-3 空间频率响应与二维傅里叶变换.....	(11)
§ 1-4 二维离散序列及其傅里叶变换.....	(16)
§ 1-5 离散傅里叶级数和离散傅里叶变换.....	(19)
§ 1-6 二维离散序列的卷积.....	(23)
§ 1-7 二维离散序列的线性变换.....	(27)
§ 1-8 二维图象函数的展开.....	(31)
1-8-1 连续图象函数展开.....	(31)
1-8-2 离散数字图象展开.....	(32)
§ 1-9 用随机过程表示图象.....	(35)
§ 1-10 三维序列简介.....	(37)
第二章 人视觉系统模型	(41)
§ 2-1 视觉系统的构造.....	(41)
§ 2-2 视觉系统对外界光刺激响应.....	(43)
§ 2-3 视觉系统的图形识别能力.....	(45)
§ 2-4 视觉系统的频率特性.....	(46)
§ 2-5 关于视觉系统等效模型的一些研究结果.....	(48)
第三章 图象数字化	(51)
§ 3-1 二维采样理论.....	(51)
3-1-1 图象空间域采样网格与空间频率域网格的关系.....	(51)
3-1-2 二维采样函数.....	(53)
3-1-3 二维采样.....	(54)
3-1-4 对图象集合采样.....	(56)
3-1-5 频率混迭问题.....	(58)
§ 3-2 正交函数展开采样.....	(59)
§ 3-3 K-L变换采样.....	(62)
§ 3-4 图象样本的量化.....	(63)
3-4-1 量化分层与图象内容的关系.....	(64)
3-4-2 量化误差及最佳量化.....	(65)
3-4-3 非均匀量化器的简单分析.....	(67)

§ 3-5 关于采样与量化的补充说明	(69)
第四章 图象信息的压缩	(73)
§ 4-1 图象空间域压缩编码	(74)
4-1-1 DPCM编码	(74)
4-1-2 增量调制 (ΔM)	(78)
4-1-3 预测编码的简单分析	(79)
4-1-4 段长编码	(81)
§ 4-2 变换域压缩技术	(82)
§ 4-3 混合压缩方法	(86)
§ 4-4 图象信息的基本概念	(88)
4-4-1. 率失真函数简介	(88)
4-4-2. 图象信息含量、多余度与压缩比	(91)
§ 4-5 图象压缩的进一步讨论	(94)
第五章 图象增强	(98)
§ 5-1 图象降质函数	(98)
§ 5-2 灰度变换技术	(104)
5-2-1 灰度直接变换技术	(104)
5-2-2 灰度直方图变换技术	(106)
§ 5-3 伪彩色 (Pseudo-Color) 增强技术	(110)
§ 5-4 微分运算增强技术	(112)
§ 5-5 图象噪声的清除	(115)
§ 5-6 图象增强综论	(118)
第六章 图象恢复技术	(120)
§ 6-1 概述	(120)
§ 6-2 几何失真和灰度失真的校正	(122)
6-2-1 几何失真的校正	(122)
6-2-2 灰度失真的校正	(123)
§ 6-3 运动模糊的等效降质函数及其校正	(124)
6-3-1 运动模糊的等效降质函数	(124)
6-3-2 运动模糊的一种校正方法	(125)
§ 6-4 空间频率域滤波法	(128)
6-4-1 逆滤波法	(128)
6-4-2 Wiener最佳滤波	(129)
6-4-3 影像匹配滤波	(133)
§ 6-5 代数解卷 —— 空间域解卷法	(133)
6-5-1 以图象最平滑为条件求最佳解	(133)
6-5-2 以图象熵极大为条件求最佳解	(135)
§ 6-6 代数解卷 —— 正交投影迭代法	(137)
6-6-1 多次正交投影迭代算法	(137)
6-6-2 投影算子迭代解法	(140)

§6-7 线性预测图象恢复算法.....	(143)
§6-8 二维同态解卷技术.....	(147)
§6-9 图象恢复其他方法综述.....	(148)

第二篇 图象分析

第七章 图象检测	(157)
§7-1 图象二值化.....	(157)
§7-2 界线(或边沿)的检测.....	(162)
§7-3 结构特性突变的检测.....	(167)
§7-4 图形匹配检测技术.....	(169)
7-4-1 模板匹配法.....	(169)
7-4-2 匹配滤波技术.....	(175)
§7-5 线条的检测与跟踪.....	(178)
§7-6 图象中物体的检测.....	(181)
7-6-1 象素灰度类聚的检测.....	(181)
7-6-2 从投影与截面估计物体.....	(182)
7-6-3 从高阶矩估计物体.....	(184)
第八章 图象几何特征分析	(186)
§8-1 邻接与连通关系.....	(186)
§8-2 边沿(界线)跟踪.....	(188)
8-2-1 并行算法.....	(188)
8-2-2 串行跟踪算法.....	(190)
§8-3 图象的几何参数测量.....	(191)
§8-4 图形骨架的提取.....	(194)
§8-5 线条的分析.....	(196)
§8-6 形状的分析.....	(202)
8-6-1 图形的几何性质.....	(203)
8-6-2 数学模型分析法.....	(205)
§8-7 纹理分析.....	(207)

第三篇 模式识别

第九章 初等模式识别理论	(215)
§9-1 引言.....	(215)
§9-2 基本数学方法.....	(216)
9-2-1 集合的一些基本概念.....	(216)
9-2-2 参量估计的基本概念.....	(217)
9-2-3 系统结构函数——模式识别中的极小熵理论.....	(218)
§9-3 基本概念和研究路线.....	(219)
§9-4 一个简单的统计模式分类器.....	(221)
§9-5 判别函数及其在特征空间的分类.....	(223)
§9-6 类别分界面和类别距离.....	(227)

§ 9-7 统计分类	(238)
§ 9-8 分类器的训练	(242)
9-8-1 确定性的训练方法	(242)
9-8-2 统计分类器的训练	(257)
9-8-3 训练结果的推广	(259)
第十章 类聚分析	(260)
§ 10-1 特征选择和特征参量简化理论	(260)
10-1-1 主轴法	(260)
10-1-2 因子分析法	(266)
10-1-3 根据类聚程度的特征排序法	(271)
§ 10-2 非监督分类的参量估计	(274)
§ 10-3 基本类聚算法	(276)
10-3-1 直接推断法	(276)
10-3-2 最大距离算法	(277)
10-3-3 结块算法	(279)
§ 10-4 类聚的数学模型和物理模型	(283)
10-4-1 类聚的数学模型	(283)
10-4-2 类聚的物理模型——吸附力模型	(285)
§ 10-5 树形类聚法	(289)
10-5-1 谱系图法	(290)
10-5-2 跨距树图表示法	(292)
10-5-3 分享共同邻域点类聚法	(294)
§ 10-6 动态类聚	(296)
第十一章 句法模式识别	(298)
§ 11-1 引言	(298)
§ 11-2 源模式及其选择	(300)
§ 11-3 模式文法	(302)
11-3-1 模式文法定义	(302)
11-3-2 模式语言的推导	(303)
11-3-3 模式文法的类型	(306)
§ 11-4 等价文法——文法的化简	(307)
§ 11-5 执行上下文无关文法的自动机	(313)
11-5-1 有限状态自动机	(313)
11-5-2 下推自动机	(317)
§ 11-6 从给定训练模式样本集推导模式文法	(320)
11-6-1 直接推导法	(321)
11-6-2 状态图综合法	(322)
§ 11-7 其他	(325)
第四篇 模式分析与应用	
第十二章 模式分析	(329)

§ 12-1	引言	(329)
§ 12-2	系统的一般结构	(329)
§ 12-3	控制模块的工作	(333)
12-3-1	抽象程序	(333)
12-3-2	迭代抽象程序	(336)
12-3-3	控制结构表示法	(338)
§ 12-4	搜索的策略	(343)
12-4-1	AND/OR搜索树	(344)
12-4-2	状态空间搜索法	(348)
§ 12-5	知识模块的结构	(349)
§ 12-6	知识的产生与获得	(355)
第十三章	图象处理与模式识别的典型应用	(359)
§ 13-1	引言	(359)
§ 13-2	图象处理基本设备	(362)
§ 13-3	本学科在医学领域中的典型应用	(364)
13-3-1	人类染色体自动分类	(365)
13-3-2	肿瘤细胞分析和分类	(368)
§ 13-4	指纹图象的计算机分析与存贮	(376)
§ 13-5	汉字图形识别	(380)
§ 13-6	结语	(382)

第一篇 图象预处理

图象预处理是图象分析的一个重要环节。它包括图象的采样、图象信息的压缩、图象信息的增强、图象的恢复等方面的内容。

第一章 理论基础

§1-1 图象的形成及其数学表示

为了利用近代计算机技术对二维图象进行处理以获得所需的信息,需要将图象用适当的数学模型来表示。为此,必须首先了解图象的形成过程以及图象与三维景物的关系,并研究二维图象的数学表示法。

图 1-1-1 表示了基本成像过程。其中,图 a 为反射光型,图 b 为透过光型。由光学的计算可知,若物体表面某点的辐照度 (Radiance) 为 $L_0(\xi, \eta, z)$ (L_0 的单位为每单位面积立体角的瓦数,记为 $\text{W}/\text{ster}\cdot\text{m}^2$), 则通过照相机透镜后,在焦距 f 的象平面

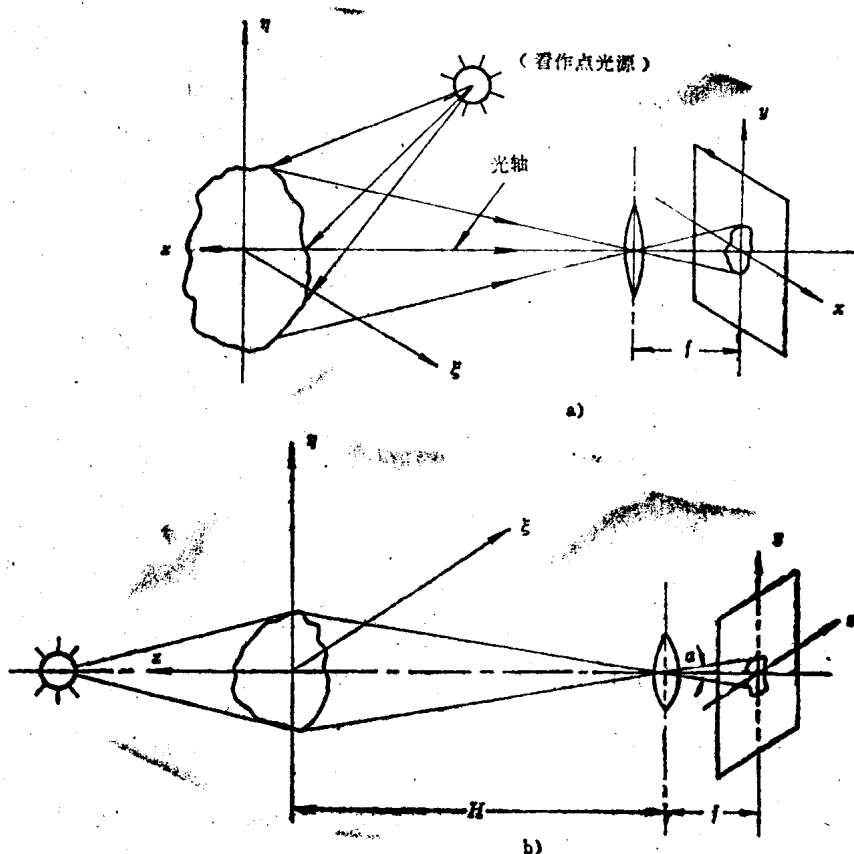


图 1-1-1 图象形成过程

处收到的照度(Irradiance)为 $E_i(x, y)$ ，单位为 W/m^2 。当 $\sqrt{\xi^2 + \eta^2} \ll H$ ，且 $\sqrt{x^2 + y^2}$ 很小时，有式

$$E_i = \pi L_0 / 4N^2(1 - M)^2 \quad (1-1-1)$$

式中， $N = f/D$ ， f 为焦距， D 为透镜直径， N 称为 f 数， $M = dA_i/dA_0$ 为放大倍数。上式是靠近光轴的情况。

若象点在象平面上偏离光轴，则应考虑 α 角的影响，并把 α 看作为在 x 、 y 方向都可以存在。此时上式修正为：

$$E_i = \pi L_0 \cos^4 \alpha / 4N^2(1 - M)^2 \quad (1-1-2)$$

上式的 E_i 与物体至透镜距离 H 无关，因为由 E_i 而引起的能量损耗很小。但 H 与像的大小却有着密切关系。

假设记录胶卷工作在线性段，则图象灰度应与胶卷密度成正比例，而密度则与 E_i 有对数关系。在动态范围不很大的情况下，可近似认为图象灰度与 E_i 成比例。所以，在上述的假设范围内，图象灰度与镜头的 f 数平方成反比，随离轴角余弦四次方值的减少而下降。故应尽可能选择在光轴附近工作。

一般地说，三维物体变成二维图象以后便失去了距离（或深度）的信息。因此，三维实物与其二维图象之间的关系问题成了研究的焦点，即怎样从图象中检测出立体信息。

假定以物体表面的法线向量为基准，设入射光线的俯仰角为 θ ，方位角为 ϕ ，反射光线的俯仰角及方位角则分别为 θ' 和 ϕ' ，见图 1-1-2，则此物体的表面元可用一个双向反射分布函数 (Bidirectional Reflectance Distribution Function) 来表征，简记为 BRDF。其定义为：

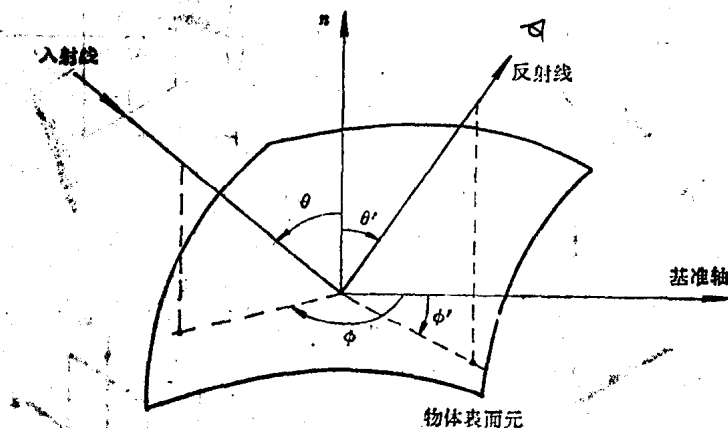


图 1-1-2 物体表面光学辐射关系

BRDF = 表面元的辐照度/表面元的照度

$$= dL(\theta, \phi, \theta', \phi') / dE(\theta, \phi) \quad (1-1-3)$$

BRDF 的单位为 $1/\text{ster}$ 。其物理意义为：若一个法线方向为 n 的表面元受到来自 (θ, ϕ) 方向的照度为 $E(\theta, \phi)$ 的光的照射，它在 (θ', ϕ') 方向反射出来的光的辐照度为 $L(\theta, \phi, \theta', \phi')$ ，则我们将在 (θ', ϕ') 方向看到的 dL 与 (θ, ϕ) 方向的 dE 之比称为

该表面元的BRDF。实际上,就一个小的表面元来说,BRDF就是该点的反射系数。由此可知,物面各点的BRDF与 θ 、 ϕ 、 θ' 、 ϕ' 有关。所以一旦计得或测知物面的BRDF函数形式,就可求出物体表面元的取向,根据一定的边界条件,即可得知物体表面形状。一般地,可用实验手段得到BRDF。所以由物体表面的辐照度便可测知物体表面的取向 $p = \partial z / \partial x$, $q = \partial z / \partial y$ 。此处的 (x, y, z) 为物体表面一点的坐标。由以上简单讨论可知,图象灰度的深浅与取向 θ 、 ϕ 、 θ' 、 ϕ' 有关。据报导,由二个从不同方向拍得的照片便可以推出物体表面形状。由上面的讨论也可知,存在着一些BRDF显著变化的特殊点,它们在图象中也表现得很特殊,这就是描述物体表面的几何性质的特征点。这些特征点构成了被研究物体的形状。

综上所述,二维图象是三维物体的投影,图象各点的灰度与物体表面相应点的取向及光源照射角度和拍摄角度有关。因此,图象灰度的大小实际上包含着第三维的信息。这是最近几年人们研究的结果。所以,在一定条件下使用各种反投影的算法,可以从图象灰度得知三维物体的形状。

以上讨论了图象的形成过程及其与实物的关系。若使用波长 (λ) 不同的光分别进行拍摄,则会得到不同的图象,这是因为被拍物体对不同波长的光有不同的响应。这种特性称为谱貌或谱特征(Spectrum signature)。现在的地球资源考察卫星和侦察卫星大都通过使用四个或七个波段对地球表面进行摄影以获得地物的全部信息。所以在计算到达镜头的光的能量时还应考虑谱特征 $V(\lambda)$ 。假设所研究的物体或景物是静止的或缓慢变化的,则图象的灰度函数为:

$$f(x, y) = \int_{\lambda_1}^{\lambda_2} V(\lambda) E(x, y, \lambda) d\lambda \quad (1-1-4)$$

由图象的形成过程知, $f(x, y)$ 应是 (x, y) 的连续函数,而且图象灰度应是正的,因为实际图象不可能表现为负灰度。这就是所谓的正性约束。故有:

$$f(x, y) \geq 0 \quad (1-1-5)$$

必须注意,这里说的正性约束只适用于原给定图象的灰度数据。处理过程中的数据有可能出现负值。

实际上,成像器件的动态范围总是有限的,所以灰度 f 也是有界的。由于实际图象总是有一定范围,故变量 x 、 y 亦是有界的。另外,因实际景物的辐照度 L 总是连续变化的,尽管这种变化可以很急剧,但总是渐变的,所以,函数 $f(x, y)$ 的导数 $\partial f / \partial x$ 、 $\partial f / \partial y$ 也是有限的。由此可得出结论:图象灰度函数 $f(x, y)$ 是一个连续的、有界的、各阶导数都存在的解析函数。这可用一个连续曲面来表示,如图1-1-3所示。

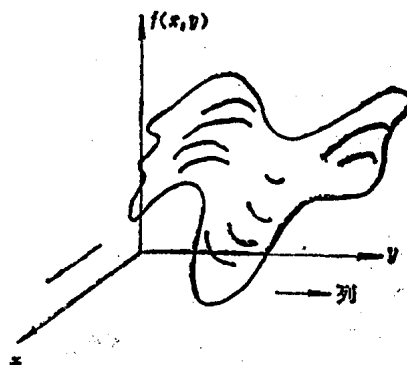


图 1-1-3 图象的灰度函数

如前所述,在遥感技术中也常用一个三维函数 $f(x, y, \lambda)$ 来表示所得的图象,其

目的是便于分析景物的谱特征。同样地, 这个三维函数也是连续的解析函数。

一般来说, 三维图象可用三维函数 $f(x, y, z)$ 来表示。如研究的是二维图象序列或运动图象, 也可用 $f[x(t), y(t)]$ 或 $f(x, y, t)$ 表示。为简单起见, 这里先讲二维函数 $f(x, y)$, 然后再将其结论推广到三维函数。

在图象分析或图象处理时, 需要将连续图象数字化, 对连续图象以适当的空间频率采样脉冲进行采样, 用得到的离散样本代替原来的连续图象。离散图象在数学上可用下面的数字灰度矩阵表示:

$$f = \begin{bmatrix} f_{0,0} & \cdots & f_{0,N-1} \\ f_{1,0} & \cdots & f_{1,N-1} \\ \cdots & \cdots & \cdots \\ f_{M-1,0} & \cdots & f_{M-1,N-1} \end{bmatrix} \quad (1-1-6)$$

上式中 MN (即 M 行 N 列) 个像素表示原来的连续函数 $f(x, y)$ 。设连续图象宽为 A , 长为 B , 则样本间的行距与列距相应为:

$$\Delta x = A/(M-1); \quad \Delta y = B/(N-1) \quad (1-1-7)$$

M 、 N 的选择将在第三章里讨论。在图象处理时可对矩阵 f 进行各种运算。

同理, 三维离散图象也可用一个三维数字矩阵表示。用 MNP 个元素表示连续函数 $f(x, y, z)$ 。

以上所述是连续图象与离散图象的最一般的数学模型。对于一些特殊图象, 则有一些相应的特殊表示方法。下面列举其中一些较典型的方法。

对于只有黑白两种灰度的二值线图 (Drawing 或 Graphics), 一般可用链码 (Chain Code) 表示。一条曲线, 不论是闭合的还是非闭合的, 均可用链码表示, 其规则是必须把曲线的斜率离散分为 8 个方向, 用 0 至 7 数字表示, 如图 1-1-4 所示。图中用三位码表示了八个离散斜率方向。对于一条给定的曲线, 可先用足够密的网格使之离散化, 然后以离曲线最近的网格点为顶点按图 1-1-4 的基准方向作折线来代替原来的连续曲线, 写出链码字符串, 标上曲线的起点坐标, 这样, 便完成了曲线的链码表示。容易看出, 链码串与折线之间是严格遵循一一对应关系的, 而折线只是连续曲线的近似表示。格子愈密, 误差愈小。图 1-1-5 所示是其中一例。

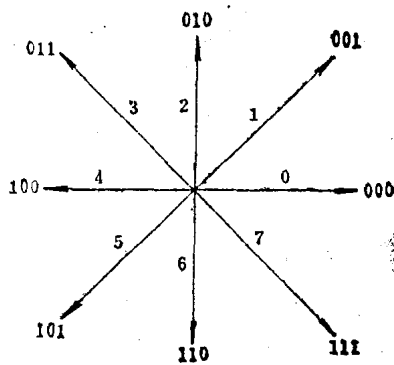


图 1-1-4 离散方向表示法

由此可知, 通过作出图象的等高线族, 并用链码表示每条等高线, 便可用许多数字串表示出原给定图象。线图的表示还可以用其他方式, 如有人使用折线各顶点的坐标序列来表示该折线, 进而用这个顶点坐标字符串来表示原曲线。

对于二值图象, 还有另外一些表示方法, 其中值得一提的是树形编码表示法。这里简单介绍四叉树 (Quad tree) 表示法。设有二值图如图 1-1-7 所示。四叉树的规则如

下：首先把整个图形分为四个子区，如图1-1-6所示。各子区的编码顺序如图中所标那样，按左上、右上、右下、左下次序编码，如灰度不均匀的块用G。表示整块灰度均匀的用W(白)或B(黑)表示。再把G分为四子块，如此下去，直到得均匀小子块为止。对

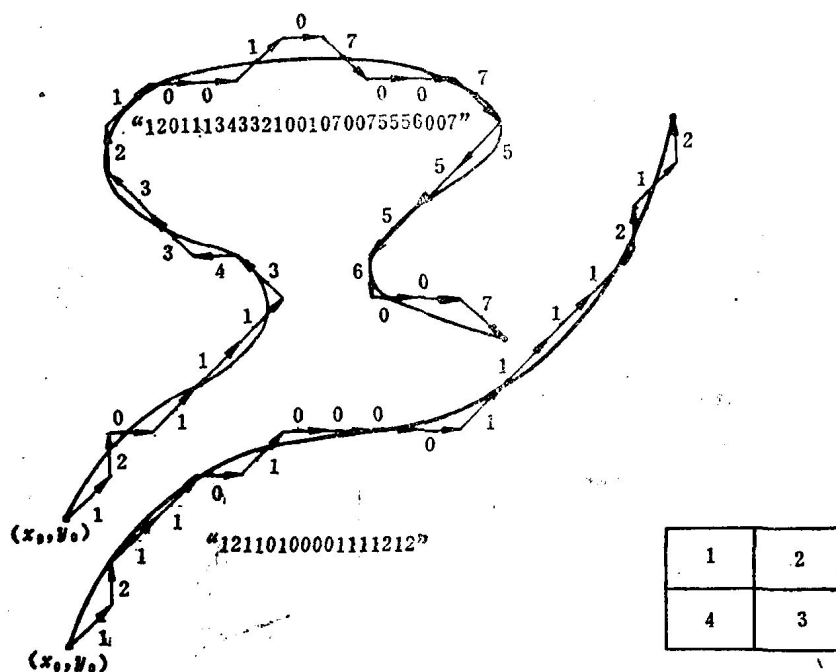


图 1-1-5 链码例

图 1-1-6 基本的图形分区编码顺序

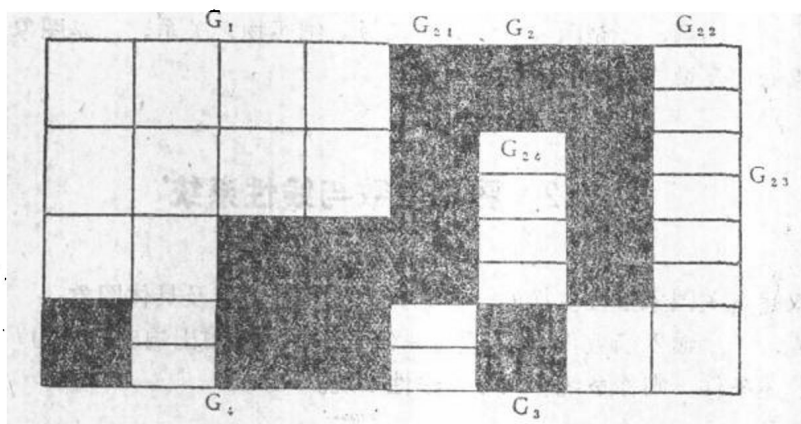


图 1-1-7 简单图形举例

于图1-1-7的例子，可首先记整幅图象为 G_0 ，由于灰度不均匀，故将整个图分为四子区，分别以 G_1 、 G_2 、 G_3 、 G_4 表示。其中 G_1 为全白，可记为W，而因 G_2 是不均匀的，应再分出四个子块，得 G_{21} 、 G_{22} 、 G_{23} 、 G_{24} 。这里， $G_{21}=B$ 。 G_{22} 不均匀，须再分四个子块，记为 G_{221} 、 G_{222} 、 G_{223} 、 G_{224} 。本例中， $G_{221}=B$ ， $G_{222}=W$ ， $G_{223}=W$ ， $G_{224}=B$ ，故得 G_{22} 的编码为BWBB。同理，将 G_{23} 分为四子区，得 G_{23} 的编码为BWBB。

仿此, 得 G_{24} 为BW WB。子区 G_2 编码完成以后, 再把 G_3 分为四子区, 得 G_{31} 、 G_{32} 、 G_{33} 、 G_{34} 。由图知 G_{31} 的编码为BW WB, G_{32} 为BW WB, G_{33} 为W, G_{34} 为WBBW。对子区 G_4 也依此处理, 得 G_{41} 、 G_{42} 、 G_{43} 、 G_{44} , 其中 $G_{41}=W$, $G_{42}=B$, $G_{43}=B$, $G_{44}=BW WB$, 最后, 得到如图1-1-8所示的整个树形图。该树图分为四层, 即0层、1层、2层、3层, 其中0层称为根, 第三层是最下层的叶。目前, 这个图形已有一些变种, 如已出现了八叉树(Oct tree)等。整个树图可以字符串的形式表示出来。

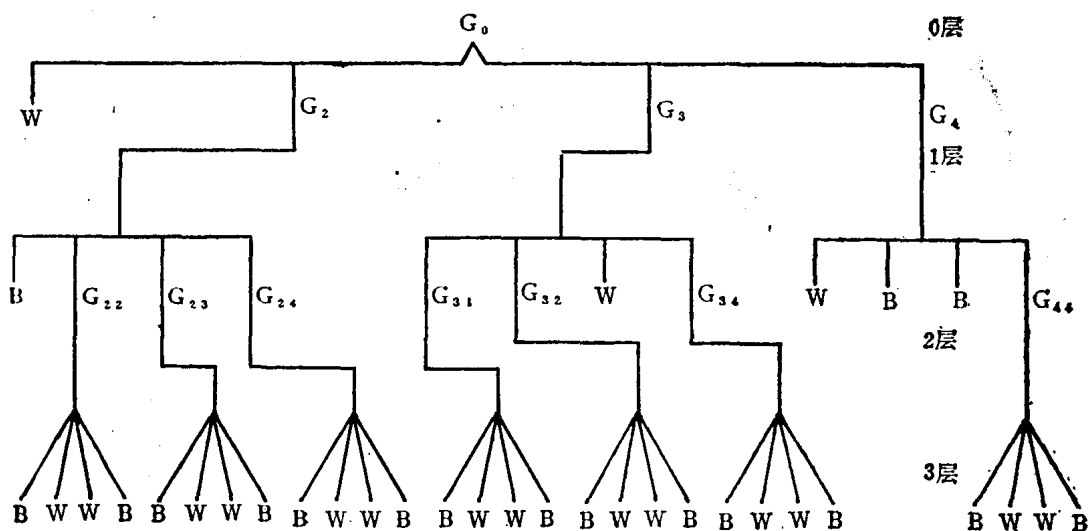


图 1-1-8 图 1-1-7 的编码树图

本节讲了两大问题, 一是图象的形成及其与三维实物之关系, 二是图象的数学表示法。这是图象处理及模式识别的最基本问题。

§ 1-2 图象函数与线性系统

本节只叙述有关图象线性变换的一般结果, 而不准备涉及具体图象。

我们知道, 一个输入函数 $f(x, y)$ 通过系统 $T[\cdot]$, 其输出响应可记为 $T[f(x, y)]$ 。如果满足下述条件, 则称系统 $T[\cdot]$ 为线性系统。

$$T[f_1 + f_2] = T[f_1] + T[f_2] \quad (1-2-1a)$$

$$T[af_1] = aT[f_1] \quad (1-2-1b)$$

由此推得

$$T\left[\iint f(x, y) dx dy\right] = \iint T[f(x, y)] dx dy \quad (1-2-2)$$

上式可用图1-2-1表示。

现引入二维矩形函数。该函数记为 $\text{Rect}(x)$ 。其定义式为(参看图1-2-2):