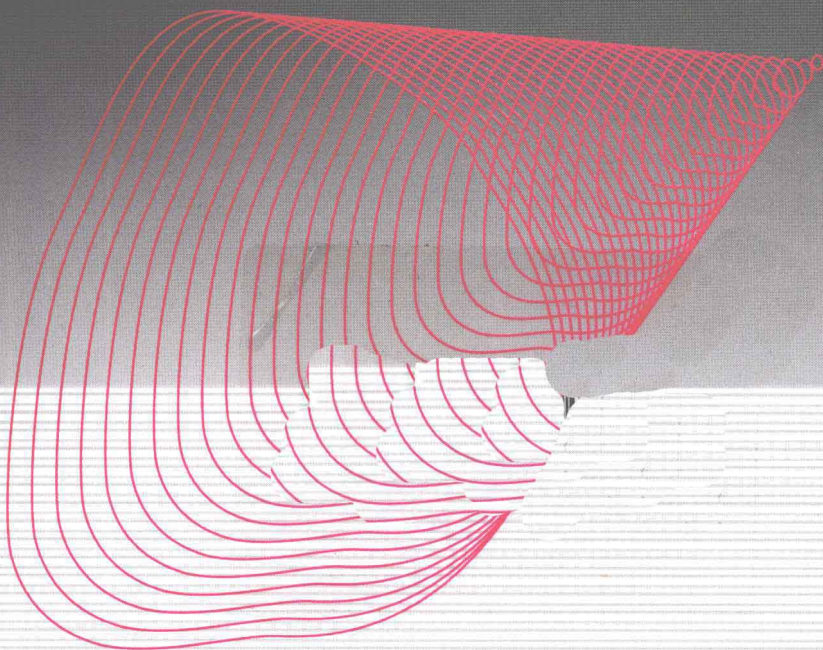


数字图像分析与模式识别

王 珣 马建峰 编著



科学出版社

数字图像分析与模式识别

王 晔 马建峰 编著

科学出版社

北 京

内 容 简 介

本书收录了近年来在图像处理、图像分析与识别领域的最新研究成果。主要内容有：图像处理与分析的基础理论与相关的Matlab仿真的编程方法；图像分析的预处理过程，包括图像滤噪、边缘检测与图像分割；图像的特征提取方法，包括基于边缘的特征提取、基于内容的特征提取、基于形状的特征提取与基于纹理结构的特征提取的典型方法；特征的选择与优化算法；图像分析的模式识别以及贝叶斯决策、线性分类、神经网络、支持向量机等分类方法。

本书给出了Matlab的仿真应用实例，可有效地帮助读者理解和掌握相关知识和技能。

本书适合于电子科学与技术、信息与通信工程、计算机科学与技术、控制科学与工程等领域的科技工作者与工程技术人员阅读，同时也可作为相关专业的研究生专业课教材与高年级本科生专业选修课教材。

图书在版编目(CIP)数据

数字图像分析与模式识别 / 王珣, 马建峰编著. —北京:
科学出版社, 2011. 9
ISBN 978-7-03-032129-9

I. ①数… II. ①王… ②马… III. ①数字图像-图像
分析 ②数字图像-模式识别 IV. ①TN911.73 ②TP391.41

中国版本图书馆 CIP 数据核字(2011)第 170473 号

责任编辑：韩卫军 王翔 / 封面设计：陈思思

科学出版社 出版

北京东黄城根北街16号

邮政编码：100717

<http://www.sciencep.com>

四川煤田地质制图印刷厂印刷

科学出版社发行 各地新华书店经销

*

2011年9月第 一 版 开本：787×1092 1/16

2011年9月第一次印刷 印张：15.5

印数：1—1 300 字数：400 千字

定价：58.00 元

前 言

图像工程是一门研究内容 and 应用范围都非常广泛的学科，随着其理论和应用的进一步拓展，图像分析与模式识别显得越来越重要，但也逐渐成为制约图像工程相关理论发展与应用拓展的主要问题之一。本书基于作者在数字图像处理和模式识别方面多年的教学经验与研究心得，以图像分析与模式识别的四个环节即图像获取、预处理、特征提取与选择、模式分类为主线，主要介绍了图像处理、图像分析与模式识别的基本概念、基本原理、典型方法与该领域的最新研究成果，并结合章节内容，介绍了 Matlab 编程实现图像处理与图像分析仿真的基本方法，给出了部分 M 函数与仿真结果。

本书主要内容分为五部分，第一部分（1—4 章），主要介绍图像处理与分析的基础理论与相关的 Matlab 仿真编程方法与部分 M 函数，并介绍了几种较新的变换技术；第二部分（第 5 章）主要介绍图像分析的预处理过程，包括图像滤噪、边缘检测与图像分割，给出了相关 Matlab 仿真程序与结果，并介绍了图像预处理的研究现状与最新方法；第三部分（第 6 章）主要介绍图像的特征提取方法，包括基于边缘的特征提取、基于内容的特征提取、基于形状的特征提取与基于纹理结构的特征提取的典型方法；第四部分（第 7 章）主要介绍特征的选择与优化算法，包括特征的预处理、PCA、ICA 等线性方法、组合优化算法（如模拟退火、遗传算法等）、基于统计假设检验的特征选择方法与 Isomap 等非线性降维方法；第五部分（第 8、9 章）主要介绍图像分析的模式识别与分类方法，有贝叶斯决策、线性分类、近邻算法、聚类算法、神经网络与支持向量机等，并给出了主要的 Matlab 仿真程序。

本书主要特点是体系完整，结构清晰。同时，本书介绍了数字图像分析与模式识别领域的最新研究成果，反映了国内外该学科的最新发展动态。书中给出了大量的 Matlab 仿真程序的编写方法与例子，不仅促进读者理解与掌握相关理论，而且可以提高读者的仿真实验能力。

本书主要读者为电子科学与技术、信息与通信工程、计算机科学与技术、控制科学与工程等领域的科学工作者与工程技术人员，同时本书也可作为相关专业的研究生专业课教材与高年级本科生专业课教材与参考书籍。

由于作者水平有限，书中错误与不当之处难免，敬请读者批评指正。

本书的出版得到了科学出版社、陕西师范大学的大力支持，并受到陕西师范大学质量工程项目资助（884217），在此表示衷心感谢。

作 者

2011 年 7 月于陕西师范大学

目 录

第 1 章 绪论	(1)
1.1 图像工程	(1)
1.2 图像工程相关的硬件	(2)
1.3 图像工程相关的软件	(7)
1.4 图像分析与模式识别	(10)
参考文献	(11)
第 2 章 数字图像表示与编码	(12)
2.1 图像数字化	(12)
2.2 图像数字化对图像信息的影响	(17)
2.3 数字图像编码基础	(19)
2.4 数字图像的预测编码	(28)
2.5 数字图像的变换编码	(31)
2.6 与数字图像表示与编码相关的 Matlab 编程	(35)
参考文献	(43)
第 3 章 数字图像的直方图与几何变换	(44)
3.1 图像直方图	(44)
3.2 基于图像直方图的图像操作	(45)
3.3 图像的几何变换	(51)
3.4 基于控制点的图像空间变换	(59)
第 4 章 数字图像的变换域处理	(64)
4.1 傅里叶变换	(64)
4.2 离散余弦变换	(70)
4.3 离散哈特利变换	(73)
4.4 方波型变换	(74)
4.5 基于特征分析的变换	(80)
4.6 小波变换	(81)
4.7 Radon 变换	(97)
4.8 Contourlet 变换	(100)
参考文献	(102)
第 5 章 数字图像的预处理技术	(103)
5.1 图像退化模型与噪声分类	(103)

5.2	图像噪声的空间域滤波方法	(109)
5.3	周期噪声与频域滤波	(121)
5.4	小波滤波	(126)
5.5	偏微分方程图像去噪	(130)
5.6	边缘检测	(133)
	参考文献	(144)
第 6 章	数字图像的特征提取	(145)
6.1	数字图像的特征与标准	(145)
6.2	基于边界的特征提取	(148)
6.3	基于区域的特征提取	(152)
6.4	基于正交矩的图像特征	(157)
6.5	基于形状的图像特征	(167)
6.6	基于纹理的图像特征	(173)
	参考文献	(177)
第 7 章	特征的选择与优化	(180)
7.1	特征选择的预处理	(180)
7.2	主成分分析 (PCA)	(181)
7.3	独立成分分析 (ICA)	(183)
7.4	基于组合优化的特征选择	(186)
7.5	基于统计假设检验的特征选择	(189)
7.6	基于流形学习的非线性降维方法	(192)
	参考文献	(198)
第 8 章	基于监督学习的模式识别	(199)
8.1	贝叶斯决策理论	(199)
8.2	线性分类器	(206)
8.3	近邻分类算法	(210)
8.4	人工神经网络	(215)
8.5	支持向量机	(221)
	参考文献	(227)
第 9 章	非监督学习的模式识别	(228)
9.1	相似性度量	(228)
9.2	聚类算法	(231)
9.3	基于核的聚类算法	(235)
9.4	基于模糊理论的聚类算法	(237)
9.5	其他聚类算法	(241)
	参考文献	(242)

第 1 章 绪 论

图像分析技术创始于 20 世纪 50 年代后期，在 20 世纪 60 年代初开始兴起，经过多年的发展，图像分析已受到许多学科的广泛重视，在机器视觉、卫星遥感、生物医学、军事、公安、交通等众多领域有广泛应用^[1]。

1.1 图像工程

1.1.1 图像工程的内容

数字图像 (Digital Image) 是模拟图像经过截取、采样、量化过程后形成的数字化表示。多媒体技术和计算机技术的发展使数字化图像信息越来越丰富。如何有效地应用这些图像信息有着非常重要的意义，而且在研究领域也受到广泛关注。图像技术的发展与数学、物理学、心理学、生理学、电子学、计算机等学科有非常紧密的联系。为了在一个整体框架下进行数字图像的研究与应用工作，1982 年，首次提出了图像工程的概念^[1]。图像工程是一门系统地研究与图像技术相关的各种理论、技术和应用的交叉学科^[2]。图像工程的内容非常丰富，根据抽象程度和研究方法的不同，分为图像处理、图像分析和图像理解三个层次，如图 1-1 所示。

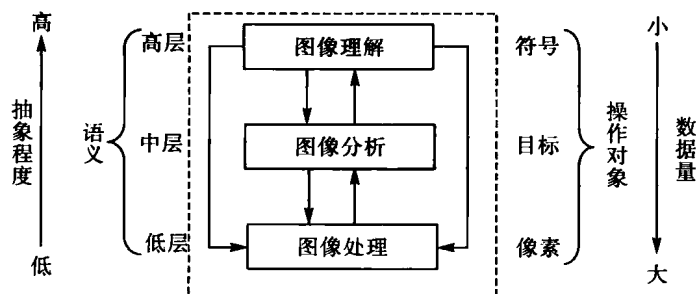


图 1-1 图像工程层次示意图

图像处理主要是在图像的像素级表示上进行操作，其重点是图像或图像表示之间的变换，其输入数据是图像的像素数据或其在某种变换域的表示方式，经过图像处理，输出结果也是图像的像素表示或某种变换域的表示方式。图像分析的主要目标是基于图像的像素级表示来产生图像或图像目标的特征表示，是图像到图像特征之间的变换。图像理解是图像特征到语义符号的变换，其输入是图像分析产生的图像特征，输出为图像内容的语义符号，如图 1-2 所示。

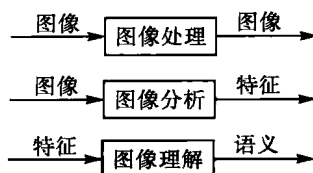


图 1-2 图像工程各层次的数据处理图示

除以上相关概念外，与数字图像相关的领域还有计算机视觉（Computer Vision）与计算机图形学（Computer Graphic）。计算机视觉研究的目的是探索与研究能够理解自然景物的方法与系统，类似于机器人的视觉系统，所以又称机器视觉（Machine Vision）。计算机图形学主要探讨用计算机将概念与数学描述的景像进行处理与生成的过程。目标是根据给定的景物描述模型、光照及成像角度生成对应的景像，如利用计算机系统为媒体进行艺术创意等。

1.1.2 与图像工程相关的学科与领域

图像工程是一门系统地研究图像理论、技术与应用的交叉学科，其研究方法与数学、物理学、心理学、生理学、电子学、计算机等有紧密的关系，研究范围与模式识别、计算机视觉、计算机图形学、信号处理等相互交叉，研究进展与人工智能、神经网络、数据挖掘、模糊逻辑等理论和技术有密切联系，其发展与应用和医学、遥感、通讯、自动化等应用领域密切相关。

图像工程与相关学科的关系如图 1-3 所示，计算机图形学主要研究基于数据描述生成图像的理论、方法与技术，模式识别会应用于图像到符号或数据特征到符号的变换，也会应用于图像理解过程中。以上学科各有侧重，但又彼此联系、互为补充。相关概念在实用中并没有严格的界线，专业或背景不同的人习惯使用不同的术语。

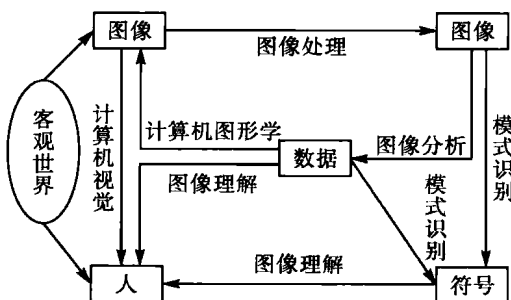


图 1-3 图像工程与相关学科的联系

1.2 图像工程相关的硬件

图像工程相关的硬件主要包括图像数字化设备、图像处理的相关硬件和图像的显示与成像设备三部分。图像数字化设备的主要任务是将自然图像转换为数字表示形式，并输入到图像处理设备中。其首先对自然图像进行采样，将其划分为若干个图像像素，然

后测量每一个像素的平均值，将其量化转换为数字形式，并将结果保存到存储中。图像数字化设备通常由采样孔、扫描机构、光传感器、量化电路和存储五部分组成，采样孔与扫描机构负责将图像划分为多个像素，光传感器负责将光信号转换为电信号，通过量化电路将其转化为数字信号，并将转化结果暂存于存储中，以用于后续的处理。

1.2.1 图像的扫描机构与光传感器

图像的采样与扫描机构主要有扫入、扫出、扫入/扫出三种实现方式。扫入方式是光源每次只照亮图像的一个像素，光传感器测量所有探测到的光强度。扫出方式是光源照亮整个场景，采样孔只允许一个像素的光线进入光传感器。在扫入/扫出系统中，场景的每个像素被一个移动的光点依次照亮，采样孔跟随移动的光点接收被照亮像素的反射光线进入光传感器。扫入/扫出方式可以有效地抑制眩光效应，主要应用于显微图像数字化过程中，但实现起来较为复杂。

光传感器根据其实现原理主要分为五类^[3]：光电倍增管、光电池、光电阻、硅传感器和结器件。光电倍增管的光电发射表面构成了半透明的光阴极（图 1-4），管壁涂以碱金属的氧化物。当具有一定能量的光子撞击充以负电荷的光阴极时，电子就从阴极表面溢出。光阴极后面是一系列带有递增正电压的倍增电极。从光阴极表面溢出的电子在电场的作用下向第一个倍增电极方向加速运动，撞击该倍增电极后会释放出更多的电子，从而产生倍增效应，这些电子在电场力的作用下向下一级倍增电极运动，同样会产生更多的电子，电子数逐级放大，最后被阳极吸收，在外电路产生的电流强度与入射光通量成正比。由于光电倍增管的倍增作用，对光照非常敏感，可以用于对弱光的检测。

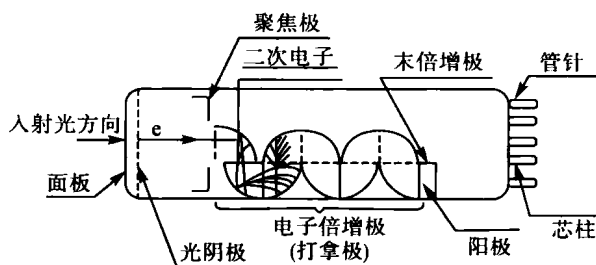


图 1-4 光电倍增管工作原理图

光电倍增管按其接收入射光的方式一般可分成端窗型（Head-on）和侧窗型（Side-on）两大类。侧窗型光电倍增管（R 系列）是从玻璃壳的侧面接收入射光，而端窗型光电倍增管（CR 系列）则从玻璃壳的顶部接收入射光。在通常情况下，侧窗型光电倍增管的单价比较便宜，在分光光度计、旋光仪和常规光度测定方面具有广泛的应用。大部分的侧窗型光电倍增管使用不透明光阴极（反射式光阴极）和环形聚焦型电子倍增极结构，这种结构能够使其在较低的工作电压下具有较高的灵敏度。端窗型光电倍增管也称顶窗型光电倍增管，它是在其入射窗的内表面上沉积了半透明的光阴极（透过式光阴极），这使其具有优于侧窗型的均匀性。端窗型光电倍增管的特点是拥有从几十平方毫米到几百平方厘米的光阴极，另外，现在还出现了针对高能物理实验用的可以广角度捕获入射光的大尺寸半球形光窗的光电倍增管。

硅传感器由高纯度硅晶体组成，在硅晶体中每个硅原子在硅晶格中与四个相邻的硅原子以共价键相连。具有一定能量的人射光子可以破坏这样一个共价键，产生一个自由电子与一个空穴。在多晶硅的表面沉积一层薄金属导电膜，加以正电压形成一个收集与保存多晶硅中释放出的自由电子的势阱。在多晶硅传感器阵列中每个势阱探测一个图像的像素。分子热运动也会随机地破坏共价键，产生与光电子无法区分的热电子，这会使硅传感器产生“暗电流”。“暗电流”的产生会对硅传感器检测结果带来噪声干扰。当然，“暗电流”的强度与温度有关，温度越高，“暗电流”越大。为减少“暗电流”的影响，常采用冷却的方法。对硅传感器进行冷却可以使其“暗电流”从室温下的每秒数千个电子减少至每秒几个电子的水平。另外过度曝光会产生过多的电子蔓延到相邻的势阱中，从而导致图像扩散性模糊。

光敏二极管是一种固态 P-N 结器件，在具有相反极性的半导体材料的结合处形成一个电场，此电场将载流子清扫出结区，产生一个阻碍电流流动的耗尽层。这样的器件通常只允许电流向一个方向流动，器件的一侧做得非常薄，以便光线穿透到结区。在工作状态下，给结区施以反向电压，这样，它只能传导很微弱的电流。在人射光的照射下，结区内部会生产电子空穴对，耗尽层中强的电场会使光照产生的大部分电子在与空穴复合之前漂移，在外电路产生电流，通过检测该电流形成光电转换，实现对入射光的检测。由于 P-N 结对反向电流具有很强的阻碍作用，所以电流的大小主要受光强的影响。

雪崩光敏二极管采用与光电倍增管类似的电子倍增效应，可以达到比普通光敏二极管更高的灵敏度。其主要原理是在耗尽层加上很高的反向电压，由光照射产生的电子被耗尽层的强电场加速，获得足以发生撞击电离的高速度，从而产生更多的电子，这种效应可导致很高的放大效应。

光敏二极管还可以工作在积分方式下，由于二极管的 P-N 结具有电容特性，可以存储反向偏置电荷，而其光导特性会使电荷与入射光能量成正比地泄漏。如果光敏二极管被定期地重新充电至某一参考电压，充电所需的电子数与入射光光能量及充电间隔时间成正比，那么在积分方式下，光敏二极管可以测量一定时间内的累积光能量。

工作在积分方式下的光敏二极管动态范围会受两个因素的限制。第一，小的结电容限制了能存储的最大初始电荷量。第二，在无光情况下出现的暗电流会使光敏二极管持续放电。这些因素会使积分周期限制在几个毫秒范围内，由于暗电流对温度敏感，降低光敏二极管的温度可以提高实际的积分时间。

光敏三极管和普通三极管相似，也有电流放大作用，只是它的集电极电流不只是受基极电路和电流控制，同时也受光辐射的控制。通常基极不引出，但一些光敏三极管的基极也被引出，用于温度补偿和附加控制等。

当具有光敏特性的 P-N 结受到光辐射时，形成光电流，由此产生的光生电流由基极进入发射极，从而在集电极回路中得到一个放大了相当于 β 倍的信号电流。不同材料制成的光敏三极管具有不同的光谱特性，与光敏二极管相比，具有很大的光电流放大作用，即很高的灵敏度。

电荷耦合器件 (CCD) 与电荷注入器件 (CID) 是基于光传感器设计的两种主要图

像检测设备。CCD 是 1969 年由美国贝尔实验室 (Bell Labs) 的维拉·博伊尔 (Willard S. Boyle) 和乔治·史密斯 (George E. Smith) 所发明的, 并因此获得了 IEEE 颁发的 Charles Stark Draper 奖章和 2009 年诺贝尔物理学奖。CCD 在摄像机、数码相机和扫描仪中应用广泛, 只不过摄像机中使用的是点阵 CCD, 即包括 x , y 两个方向用于摄取平面图像, 而扫描仪中使用的是线性 CCD, 它只有 x 一个方向, y 方向扫描由扫描仪的机械装置来完成。

CCD 是由基片上的光传感器阵列形成一个图像像素检测阵列, 每个像素产生的光电子被收容到对应的势阱中, 势阱中的光电子主要有三种读出结构: 全帧结构、行间传送结构与帧传送结构。全帧 CCD 曝光后在读出过程中必须保持黑暗, 然后将 CCD 最底行的电荷移出至外电路, 一次只读一个像素。当底行被移空后, 所有行的电荷向下移动一行, 重复此过程直至读完最后一行。行间传送结构 CCD 中, 传感器的偶数列被一个不透明的掩膜覆盖, 只有奇数列曝光, 曝光后每个奇数列势阱中的电荷被移动至相邻的偶数列中, 读电路只读偶数列中的电荷数量。帧传送结构 CCD 采用双倍高度的传感器阵列, 上部传感器阵列用来曝光, 下部的传感器阵列用不透明的掩膜覆盖用来存储, 曝光结束后, 上部传感器阵列中的电荷被快速移动到下部对应行中, 被外电路读取。帧传送结构 CCD 可以同时采集图像与读取图像, 可以获得高的图像探测速度, 主要用于视频图像的采集。

CCD 由于其对光的响应呈现出良好的线性特性, 而且几何畸变很小等优点被广泛应用于各种图像传感设备中。既可以应用于视频图像的采集, 也可以通过长时间的累积曝光来捕捉较弱的图像信息, 如应用于天文图像获取或荧光显微镜等设备中。由于硅晶体中存在的暗电流所带来的干扰, 对于长时间曝光的 CCD, 在应用中应对其进行降温冷却, 以降低暗电流的影响。通常 CCD 采集图像的过程中除暗电流的影响外, 还存在两种可能的噪声干扰, 分别被称为读出噪声与光子噪声。读出噪声是一种片内电路产生的随机噪声, 与芯片设计的技术相关。在快速读取过程中, 读出噪声的影响较为显著, 在这种情况下, 暗电流与光子噪声成分较小, 读出噪声成为噪声干扰的决定因素。光子噪声是由光的量子性导致的, CCD 实际接收的光子数服从泊松分布, 分布方差近似等于光照均值的平方根, 在强曝光情况下成为噪声的主要因素, 另外过度曝光会使过剩电子蔓延到相邻像素, 导致图像扩散性模糊。

CID 与 CCD 的主要区别是它们的读出方式显著不同。在每个像素对应的硅晶体两侧, 各有一个被金属氧化物与硅晶体隔离的电极, 其中一个电极与同一列所有像素对应的电极互连, 另一个电极与同一行所有像素对应的电极互连, 这样可以通过控制其行列地址来定位某个硅晶体。

控制 CID 的两个电极的电压可以使 CID 工作在图像积分、破坏性读与非破坏性读三种模式下。第一种模式是给两个电极加以同样的电压, 对于 N-型硅片加正电压, 如果是 P-型硅片则加负电压。光电子在两个电极的控制下积累下来, 实现对图像的积分。如果将两个电极电压降至零伏或更低, 积累的光电子会注入下面的基片, 同时在外电路引起电流变化, 电流的大小反映了被积累的光电子数目, 实现图像读出。然而, 读出图像的同时, 势阱中的载流子全部消失, 所以是一种破坏性读的过程。当一个电极保持电

压不变, 另一个电极的电压降至零伏或更低, 积累的载流子会向有保持电压的电极移动, 同时在外电路引起电流变化, 电流的大小反映了被积累的光电子数目, 实现图像读出。但与破坏性读不同的是, 载流子仍然保持在势阱中, 可以通过控制两极电压变化, 实现多次的重复读, 因此被称为非破坏性读出方式。由于可以通过地址选择控制部分或全部势阱, CID 既可以读出全部图像, 也可以只读出图像的部分子图。

与 CCD 比较, CID 主要在以下方面性能优于 CCD。首先, 在相邻势阱中, CID 不存在如 CCD 那样的内建通道相连, 因而由于曝光过量引起的过剩电子只会被本势阱的基片吸收, 不会像 CCD 一样蔓延到相邻势阱中, 所以 CID 对于扩散模糊与辐射危害与 CCD 相比有更高的鲁棒性。另外, 读出程序可以利用非破坏性读监控势阱的载流子产生情况, 并释放那些在图像累积周期结束前填满的单个势阱中的载流子, 实现对扩散模糊与辐射危害进一步控制。其次, 由于 CID 不在阵列中移动载流子, 晶体点阵的缺陷只影响单个像素, 不会像 CCD 一样影响整行像素的读取。最后, CID 可以进行部分读和非破坏性读, 但是 CID 与 CCD 相比, 光检测的灵敏度较低, 所以主要应用在需要其随机访问能力、非破坏性读和抗扩散模糊特性的特定场合。

1.2.2 图像的显示与输出

图像显示是数字图像处理的最后一个环节, 然而对于数字图像分析来说, 图像显示并不一定很重要, 因为图像分析是以数据或语义的形式给出结果的。图像显示除考虑图像信息本身以外, 还要与人眼视觉特性相匹配。一般来说, 人眼对灰度级不十分敏感, 只能分辨约 40 级的灰度, 但对由灰度级差别所形成的边缘比较敏感。相对来说, 人眼对彩色信息较为敏感, 对图像中的细节与边缘有较强的分辨能力。当然, 图像显示系统在图像工程中也是十分重要的, 一幅经过数字图像处理方法优化的图像很可能在最后一个环节被一个有噪声或不精确的显示系统劣化。

图像显示主要有永久性和暂时性两种形式, 永久性显示是通过永久性改变记录介质的光吸收特性, 在纸、胶片或其他媒介上形成图像的硬拷贝。暂时性显示是在显示屏上产生一幅暂时的图像。对于图像显示系统来说, 主要影响其显示性能的因素有图像分辨率、光度分辨率、线性特性、低频响应及噪声特性等五个方面。

图像分辨率一般定义为单位长度像素点的数量, 通常以每英寸上的像素点数作为度量 (PPI, Pixels Per Inch), 对于图像输出来讲, 通常以每英寸长度上的点数 (DPI, Dots Per Inch) 作为输出图像的分辨率度量。PPI 或 DPI 值越高, 图像细节越清晰。

光度分辨率是指图像输出设备在每个像素位置产生亮度或光密度的精度, 即图像输出设备能产生的离散灰度级的总数, 其主要取决于控制每个像素的比特数和输出设备的设计技术。如果控制每个像素的比特数为 k , 则理论上可以产生的灰度级总数为 2^k , 但实际的灰度级总数由于受输出设备设计技术的限制可能会低于理论灰度级数。另一个影响有效灰度级数的因素是输出设备的内部噪声, 由于该噪声的存在, 有效灰度级的数量会进一步减少。

另一个重要的显示特性是灰度的线性特性。任何图像显示设备都有一条输入灰度与输出亮度的变换曲线, 理论上这条曲线应该是线性的, 而且保持不变。但是实际上一般都会有一定的偏离。幸运的是, 人眼并不是很精确的光度计。轻微的非线性偏离很难被

觉察。但如果变换曲线的两端有明显的非线性,则在亮区或暗区可能会丢失信息造成图像质量下降。

低频响应主要说明图像显示系统显示大块等灰度区域的能力,这种能力主要依赖于显示点的形状、点间距、显示系统的幅度噪声与位置噪声。理想情况下大块等灰度区域应以均匀一致的亮度显示出来。一般来说,不同的显示系统像素的极性定义是不同的,但无论极性如何定义,零亮度区域总会呈现均匀一致。因而低频响应的性能对低亮度区域影响不大,大块等灰度区域的均匀性只有在显示中等及其以上亮度时才会成为问题。

均匀性主要取决于相邻像素能否很好的拼贴起来,某些图像输出设备利用矩形孔成像,产生互相精确贴合的矩形像素,因此能获得极好的均匀性,但作为数字图像显示常用设备的 CRT 是利用圆形光点来形成图像的,相邻像素的拼贴会影响大块等灰度区域的均匀性,为了保证好的低频响应,圆形光点应保持一定的间距,但是较大的点间距会影响显示系统的高频响应。因此,显示系统的低频响应与高频响应给点间距的选择带来了互相矛盾的制约。在实际应用中点间距的选择往往是两者的折衷。

显示的噪声特性会引起显示点亮度与位置两方面的变化,主要有幅值噪声、点位置噪声两种表现形式。幅值噪声主要由显示系统亮度通道的随机干扰所导致,点位置噪声是由于偏转或扫描电路的不精确或随机干扰所导致。一般的点位置噪声不会给图像带来可觉察的几何畸变,但是位置噪声会放大点之间的相互影响,所以要获得好的显示效果必须很好地控制点位置噪声。

1.2.3 数字图像处理的硬件

由于速度与储存容量的要求,早期的数字图像处理系统主要采用小型机与工作站,随着计算机性能的快速提高,以 PC 机为主的图像处理系统得到了广泛应用。由于数字图像处理系统的微型化要求,数字信号处理器(DSP, Digital Signal Processor)与 FPGA 技术在数字图像处理系统中的应用日趋广泛。与通用的微处理器相比,数字信号处理器的功能在许多方面都作了扩充和增强,在结构上采用了哈佛总线结构,在硬件上获得系统设计的极大灵活性。DSP 芯片的可编程性和强大的处理能力使其在图像处理领域得到广泛的应用和发展。

1.3 图像工程相关的软件

数字图像工程中与软件相关的需求主要有两个方面,在研究探索与算法设计阶段,需要使用方便、功能强大的仿真平台;在工程实现阶段,需要执行效率高、可移植性好的软件开发工具。常用的数字图像处理仿真平台是由美国 MathWorks 公司推出的 Matlab, Matlab 是 Matrix Laboratory 的缩写,具有强大的矩阵运算和处理功能,其图像处理工具箱提供了丰富的图像处理函数,从而大大节省了编写低层代码的时间。但是 Matlab 属于解释方式的编程语言,相对于编译方式的编程语言执行速度慢,不能产生脱离 Matlab 环境的可执行二进制代码,所以比较适合作为图像工程的研究探索与算法设计阶段的仿真测试平台。为了将 Matlab 的强大功能应用于图像工程实际应用中,Mathtools 公司的 Matcom 工具可以把 Matlab 的 M 文件翻译为 C 或 C++ 语言,但其翻

译的代码不够简洁,而且功能有限。因此,一般在工程实践中,往往直接用编译方式的编程语言进行设计,常用的设计工具是 Microsoft 公司的 VC++。VC++ 是一种具有高度综合性能的编译型软件开发工具,用其开发出来的程序具有运行速度快、可移植性好的特点。

1.3.1 Matlab 简介

Matlab 是一种解释型的高性能语言,其基本数据元素为并不要求确定维数的数组,这就允许编程者用公式化方法求解很多与矩阵相关的计算问题。在 Matlab 的环境中,集成了计算、可视化与编程,如图 1-5 所示,其用户界面由命令窗口、工作空间浏览窗口、当前目录窗口、历史命令窗口和一个或多个图形窗口五个子窗口构成。其中,图像窗口只有在用户用 `imshow` 指令显示图像时出现。

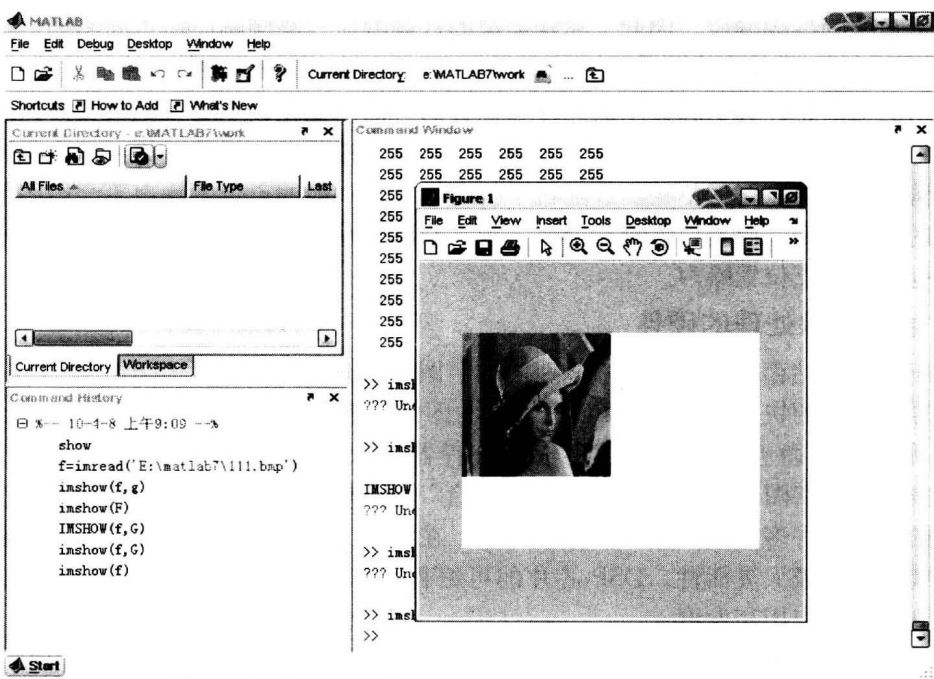


图 1-5 Matlab 界面图示

Matlab 的命令窗口是用户在 Matlab 提示符 (`>>`) 处输入 Matlab 命令或表达式的地方,同时这些命令的输出也会显示在命令窗口。工作空间浏览窗口主要显示用户在使用 Matlab 过程中创建的变量集及相关信息,这些变量集构成了 Matlab 的工作空间。在工作空间浏览窗口双击某变量会启动 Matlab 数组编辑器,它可以用来得到相关变量的详细信息,并在特定情况下编辑这些变量的相关内容。在 Matlab 中也可以用 MAT 文件保存工作空间中的变量,如果要保存工作空间浏览窗口中的所有变量,在该窗口内的任何空白处单击鼠标右键,在弹出的菜单中选择 `Save Workspace AS`,选择保存路径即可保存工作空间中的变量。如果要保存工作空间浏览窗口中部分变量,选中这些变量后,单击鼠标右键,在弹出的菜单中选择 `Save Workspace AS`。形成的文件都以双精度

二进制格式保存,扩展名为.mat。重新调入保存的工作空间变量文件主要有两种方式,第一种方式是在工作空间浏览窗口的工具条上点击文件夹按钮,另一种方式是在命令窗口或者在程序中使用 load 或 save 命令实现工作空间的保存与恢复。

当前目录窗口显示当前的工作路径,在当前目录窗口中点击右边的箭头可以显示最近使用过的路径列表。点击箭头右边的按钮可以改变工作目录。任何在 Matlab 运行的文件必须保存在工作目录或 Matlab 的搜索路径中,添加或更改 Matlab 的搜索路径方法是在 Matlab 的 File 菜单中选择执行 Set Path 命令。为了使用方便,通常应该将常用的目录添加到 Matlab 的搜索路径中。

历史命令窗口显示用户在本次启动 Matlab 后用过的命令列表,如果要重复执行某命令,可以在历史命令窗口双击该命令,也可选择一组命令序列,点击鼠标右键,在弹出菜单中选择相应操作。

如果在程序或命令中使用 imshow() 函数,Matlab 将会弹出一图像显示窗口,将对应图像显示在该窗口中,Matlab 可以根据用户要求弹出多个图像显示窗口用以显示多幅图像。Matlab 的可执行程序代码通常保存在 M 文件中,如果要编辑或创建一个 M 文件,使用 edit 命令,Matlab 会弹出一个文本编辑窗口,接收用户输入的指令序列形成 Matlab 程序。如果要修改已经创建的 M 文件,可以在命令窗口中键入 edit 文件名,系统将会打开该文件供用户修改编辑。当然也可以在历史命令窗口中创建 M 文件,在历史命令窗口选择要创建为 M 文件的命令序列,点击鼠标右键,在弹出菜单中选择 CREATE M FILE 菜单项,可以将历史命令窗口中的命令序列创建为 M 文件。

1.3.2 VC++简介

VC++是 Microsoft 公司推出的开发 Win32 应用程序面向对象的编译型语言,其提供的 Microsoft 基础类库(MFC, Microsoft Foundation Class Library)集成了操作系统提供的 Win32 应用编程接口(Win32 API, Application Programming Interface)可以支持完全面向对象的方法来设计 Win32 应用程序。另外,VC++具有很好的操作系统平台支持和可移植性,形成的可执行代码有较高的执行效率,被广泛应用于图像工程应用系统开发中。

VC++的 Developer Studio 包含了编写程序源代码的文本编辑器、设计用户界面的资源编辑器、建立项目配置的项目管理器、调试程序的集成调试器等工具,同时还提供了应用程序向导工具(App Wizard)和类向导工具(Class Wizard)。应用程序向导工具可以生成各种不同类型的具有 Windows 界面风格的应用程序的基本框架,然后使用类向导工具完成创建新类、定义消息处理函数、重载虚拟函数等操作。

VC++在应用开发中主要涉及两个基本概念,第一是项目(Project),开发一个应用就是生成一个项目。当第一次创建一个应用时,应选择创建一个项目,此时 VC++会自动创建一个工作区,并把新建的项目在该工作区打开,以后对该项目进行修改、补充等工作,只要打开对应的项目或工作区即可。第二个基本概念是工作区(Workspace),VC++以项目工作区的方式来组织文件、管理项目与项目配置,通过项目工作区可以访问或修改对应项目的所有元素。首次创建项目工作区,VC++会创建三个文件,一个项目工作区文件、一个项目文件和一个工作区设置文件。项目工作区文

件用以描述工作区及其内容,扩展名为 .dsw。项目文件用以记录项目中各种文件的名称与路径,扩展名为 .dsp。工作区设置文件主要记录工作区属性设置,扩展名为 .opt。

创建或打开项目工作区,项目工作区窗口中显示与项目有关的内容,项目工作区窗口主要有三个面板:File View, Resouce View, Class View, 分别用以显示项目中定义或创建的文件、资源及 C++ 的类。通过这三个面板,可以定位到项目中的任一位置,对其进行相应的访问与编辑。

数字图像文件格式很多,如果由用户编写各种图像文件的读写、编码、解码等程序,用户必须对各种图像文件的细节、编码与解码有非常详细的了解。为了减轻用户的负担,可以使用 imageload.dll 动态链接库很方便地实现常见图像文件的读写。imageload.dll 动态链接库支持 BMP, CIF, JPG, PCX, TGA, TIF 6 种格式的图像文件读写功能,提供了 22 个函数,其中 18 个函数与数字图像文件的读写有关,这些函数分为三类:Load 函数用于读出图像文件;Save 函数用于保存图像文件;Get 函数用于获取图像文件的相关信息。

1.4 图像分析与模式识别

图像工程是一门研究内容 and 应用范围都非常广泛的学科,随着其理论和应用的进一步拓展,图像分析与模式识别显得越来越重要,但也逐渐成为制约图像工程相关理论发展与应用拓展的主要问题之一。

图像识别属于模式识别中的统计模式识别问题,主要包括图像获取、预处理、特征提取与选择、模式分类四个步骤(图 1-6)。图像预处理的主要目的是提高图像的质量、强化某方面的特征或者将待识别的目标分离出来,主要涉及图像恢复、图像增强与图像分割技术。图像恢复的目的是为了提高图像的质量,如去除噪声、提高图像清晰度等。图像增强不考虑图像降质的原因,突出图像中所感兴趣的部分,如强化图像中的高频分量,可使图像中物体的轮廓清晰,细节明显;强化低频分量,可减少图像中的高频噪声影响等。图像分割是一种重要的图像技术,在对图像的研究和应用中,往往仅对图像中的某些部分感兴趣,它们一般对应图像中特定的、具有独特性质的区域。为了辨识和分析目标,需要将它们分离提取出来。



图 1-6 图像识别过程框图

基于图像像素级的表示很难实现图像识别,所以经过预处理的图像应该在像素级表示的基础上进一步提取其主要特征,特征提取从数学上来讲是从高维像素空间到低维特征空间的变换,但是如何提取能够很好的表示图像内容而且有利于模式分类的低维特征是图像识别研究的难点与重点之一。特征选择是对提取的低维特征进行进一步选择与优化,使选择后的特征更有利于模式分类。模式分类过程是基于特征提取与选择过程所形成的图像特征,应用模式识别的方法对图像进行分析与识别。特征选择、模式分类与模式识别领域的发展有非常大的关联。

参 考 文 献

- [1] Fu K S, Kunii T L. Picture Engineering. Springer-Verlag, 1982.
- [2] 章毓晋. 图像工程. 北京: 清华大学出版社, 1992.
- [3] K. R. Castleman. 数字图像处理. 朱志刚等译. 北京: 电子工业出版社, 2002.