

图像特征提取与语义分析

赵捷 著



重庆大学出版社

图像特征提取与语义分析

赵 婕 著

重庆大学出版社

内容提要

图像特征提取与语义分析是通过提取图像底层视觉特征,然后利用图像分析技术实现图像内容的语义描述、分类和理解。其核心是确定图像底层特征与高层语义之间的映射关系,这正是计算机视觉领域当前研究的热点与难点。本书从基本概念入手,系统地介绍了图像分析的基本理论与方法,涉及数字图像处理的基础知识、特征提取与图像表示、图像分割与目标识别、场景理解等相关问题。并在书中加入了应用实例以及实验结果图片,突出了理论和实践相结合的特点。全书深入浅出、图文并茂、文字描述简单易懂。

本书可作为高等学校计算机专业和其他信息类专业研究生和高年级本科的参考书目,也可供从事计算机视觉等相关领域研究的科研技术人员阅读参考。

图书在版编目(CIP)数据

图像特征提取与语义分析/赵捷著. —重庆:重庆大学出版社,2015.6

ISBN 978-7-5624-9150-7

I. ①图… II. ①赵… III. ①图象数据库—情报检索—IV. ①G354.49

中国版本图书馆 CIP 数据核字(2015)第 128849 号

图像特征提取与语义分析

赵 捷 著

策划编辑:周 立

责任编辑:周 立 版式设计:周 立
责任校对:关德强 责任印制:赵 晟

*

重庆大学出版社出版发行

出版人:邓晓益

社址:重庆市沙坪坝区大学城西路 21 号

邮编:401331

电话:(023) 88617190 88617185(中小学)

传真:(023) 88617186 88617166

网址: <http://www.cqup.com.cn>

邮箱: fxk@cqup.com.cn(营销中心)

全国新华书店经销

POD:重庆新生代彩印技术有限公司

*

开本:787×1092 1/16 印张:13 字数:316千

2015年6月第1版 2015年6月第1次印刷

ISBN 978-7-5624-9150-7 定价:30.00元

本书如有印刷、装订等质量问题,本社负责调换
版权所有,请勿擅自翻印和用本书
制作各类出版物及配套用书,违者必究

前言

人类通过视觉可以直观地感知和理解现实世界,而图像是视觉信息的主要载体,图像信息具有直观、形象、易懂和信息量大等特点,因此,图像成为人类获取和交换信息的重要来源。随着计算机技术和互联网的发展,图像技术已经在航空航天、生物医学、工业检测、机器人视觉、军事制导、文化艺术等应用领域收到广泛重视并取得了重大的开拓性成就。

图像的底层视觉特征是对图像进行语义分析的基础和前提,然而图像底层视觉特征和用户的高层语义之间存在“语义鸿沟”,这给图像数据的分析、描述、分类和理解带来了巨大的挑战。为了解决语义鸿沟问题,图像语义理解逐渐成为计算机视觉领域研究的热点问题,如何使计算机具有类似于人类识别图像的思维推理能力,一直以来都是研究的重点与难点。近年来,统计和机器学习理论与方法作为人工智能领域的前沿,已成为图像语义分析研究发展的重要推动力。

本书从基本概念入手,系统地介绍了图像分析的基本理论与方法,主要内容包括数字图像基础知识、特征提取与图像表示、图像分割方法及实现、目标识别、基于语义特征的场景理解、基于上下文信息的语义分析。为了使读者更容易理解图像分析的方法和理论,本书使用理论与实践相结合的方法,从基本概念出发,并加入实验实例,由浅入深地介绍了图像特征提取与语义分析的基本理论与经典方法。

在本书的编写过程中,参考和引用了国内外出版的大量文献及网站资料(这些资料在本书的参考文献中尽量列出,若有遗漏深表歉意),同时,也融入了作者的研究工作和经验。在此,对本书所引用的文献的作者深表谢意!另外,本书的编写得到了太原学院各位领导的支持和帮助,在此表示感谢!

近年来,作者一直从事计算机视觉、模式识别领域的研究工作,积累了一定的心得体会。但是图像理解问题博大精深,而作者的水平有限,书中可能存在不足或不妥的地方,恳请读者给予批评和指正。本书的读者对象主要是那些进行图像理解技术研究和产品开发的技术人员。同时,本书也可以作为计算机视觉、模型识别领域的入门参考书。

作者
2015年3月

目 录

第 1 章 数字图像基础知识	1
1.1 概 述	1
1.2 图像采集与处理	7
1.3 颜色视觉和色度图	25
1.4 像素的邻域和连通性	31
1.5 频域变换方法	32
第 2 章 特征提取与图像表示	39
2.1 基本运算	40
2.2 有用信息的增强	45
2.3 颜色特征描述	59
2.4 纹理特征分析	61
2.5 形状表示与描述	64
2.6 图像表示	76
第 3 章 图像分割方法及实现	80
3.1 灰度阈值分割法	80
3.2 边缘检测算子	86
3.3 方向算子	91
3.4 霍夫变换	94
3.5 区域分割	96
3.6 基于图论的分割	100
3.7 基于聚类的分割	104
第 4 章 目标识别	111
4.1 图像匹配	111
4.2 场景与目标的关联性	115
4.3 识别中的优化技术	120
4.4 Boosting 识别方法	124
第 5 章 基于语义特征的场景理解	129
5.1 图像理解控制策略	130
5.2 高层语义特征	132
5.3 基于分类器的图像标注	139

第6章 基于上下文信息的语义分析	163
6.1 上下文信息	164
6.2 上下文关系模型	167
6.3 图像语义分析	172
参考文献	192

第 1 章

数字图像基础知识

随着人类社会的进步和科学技术的发展,人们对信息处理和信息交流的要求越来越高。图像信息具有直观、形象、易懂和信息量大等特点,因此,它是人们日常生活中接触最多的信息种类之一。在现代社会生活中人类已经离不开图像。另外,随着计算机技术和互联网的发展,图像技术已经渗透到各个科技领域。近年来,图像信息处理已经得到一定的发展,但随着对图像处理要求的不断提高,应用领域不断扩大,图像理论也在不断提高、补充和发展。图像的处理已经从可见光谱扩展到红外、紫外等非可见光谱,从静止图像发展到运动图像,从物体的外部延伸到物体的内部,以及进行人工智能化的图像处理等。

所谓“图”是物体投射或反射光的分布,“像”是人的视觉系统对图的接收在大脑中形成的印象或反映。因此,图像是客观和主观的结合。图像是客观对象的一种相似性的、生动性的描述或写真,是人类社会活动中最主要的信息源。或者说图像是客观对象的一种表示,它包含了被描述对象的有关信息。它是人们最主要的信息源。据统计,一个人获取的信息大约有 75% 来自视觉。俗话说“百闻不如一见”“一目了然”便是非常形象的例子,都反映了图像在信息传递中的独特效果。如何使计算机具有类似于人类识别图像的思维推理能力,一直以来都是计算机视觉领域研究的重点与难点。图像作为研究的基本对象,本章主要介绍图像处理、分析与理解中所包含的基本概念、性质与技术,是后续特征提取与语义分析内容的基础知识。

1.1 概 述

图像从视觉特点,分为可见图像和不可见图像,如图 1.1 所示。其中,可见图像中包括照片、几何线条图、画和光学器件产生的图像。不可见的图像包括不可见光成像,如红外成像、微波成像、X 光成像、超声波成像、 γ 射线成像和数学物理模型生成的图像,如温度、压力等的分布变化图。

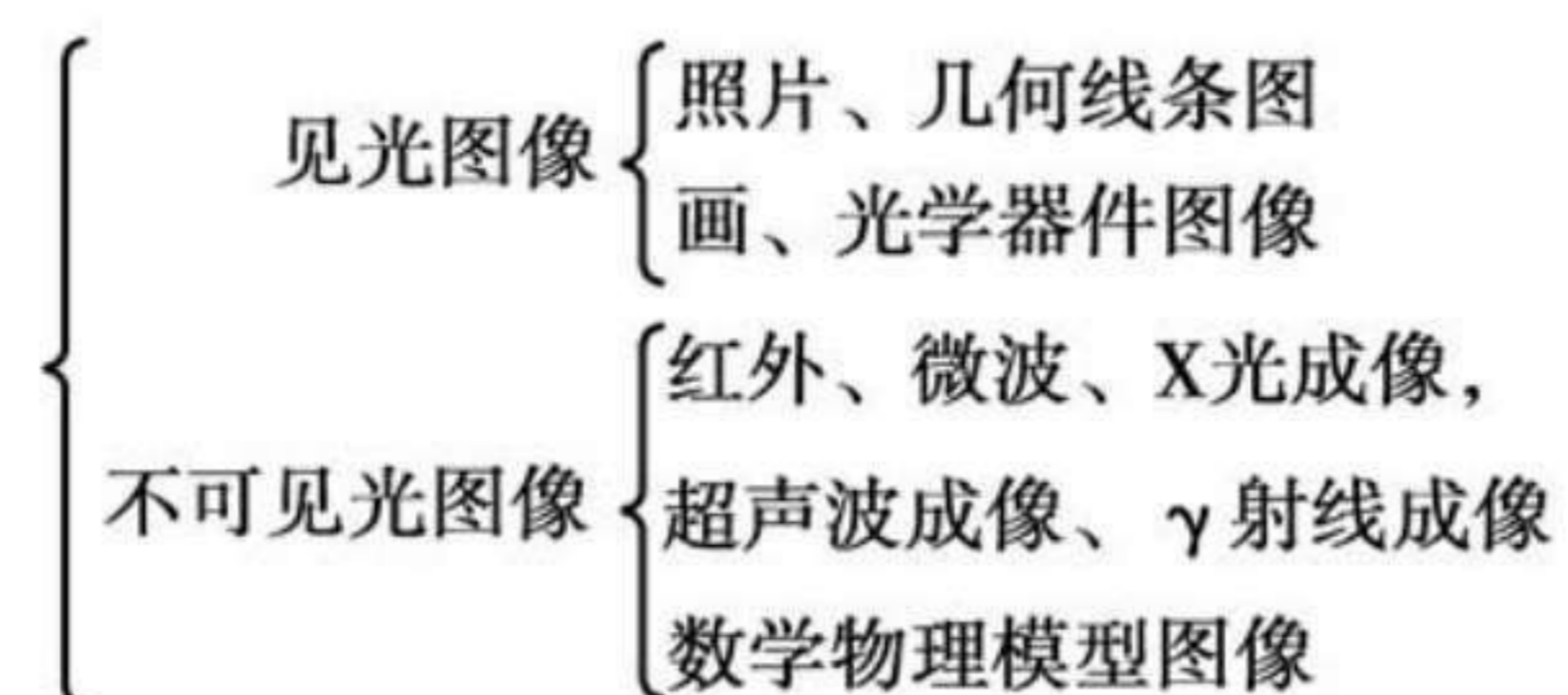


图 1.1 图像的种类

图像从图像空间坐标和明暗程度的连续性,可分为模拟图像和数字图像。模拟图像:图像是连续的,即用函数 $f(x,y)$ 表示的图像。其中 x,y :表示空间坐标点的位置, f :表示图像在点 (x,y) 的某种性质的数值,如亮度、灰度,色度等。 $f(x,y)$ 可以是任意实数,用连续函数来描述。图像的坐标系统如图 1.2 所示。

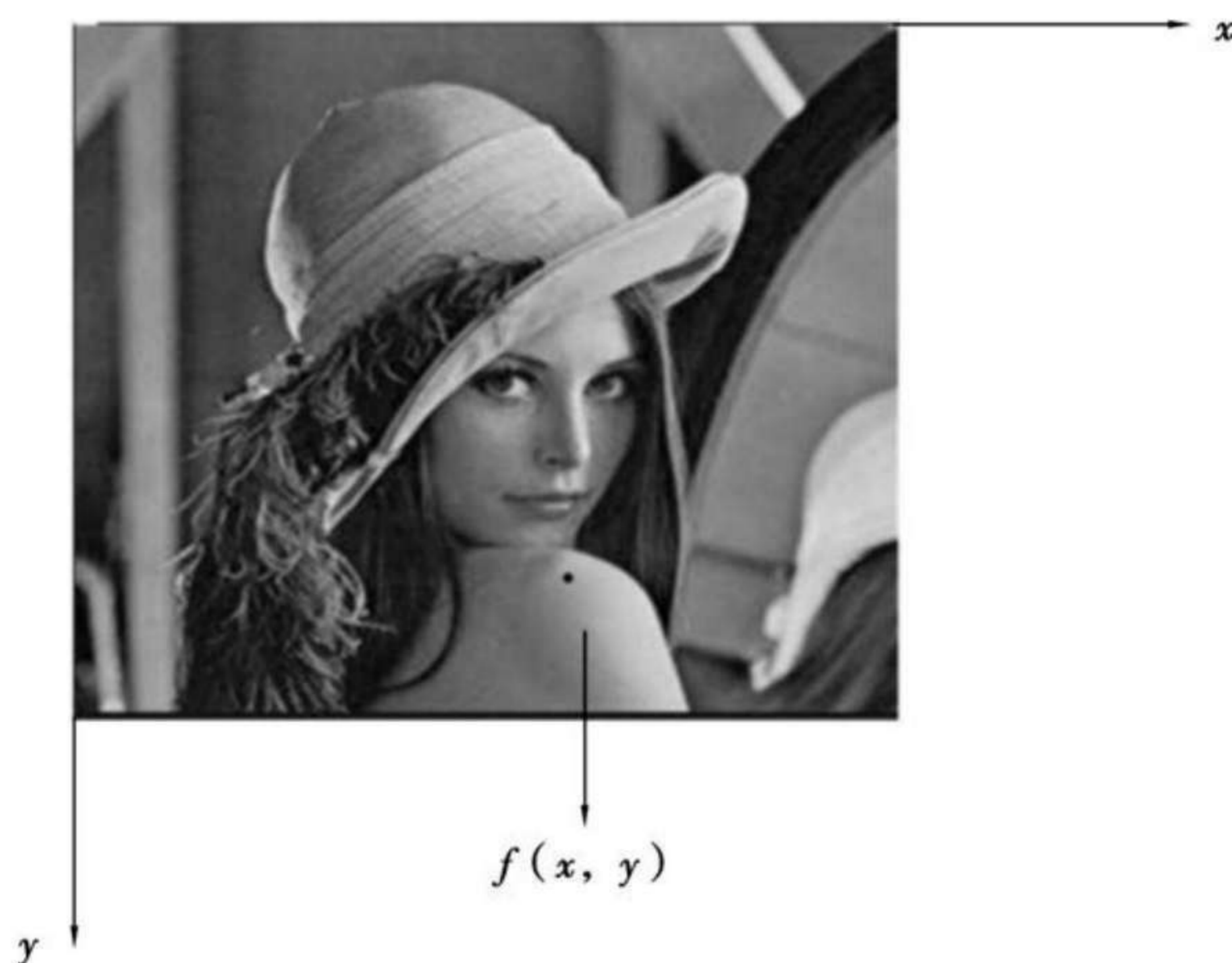


图 1.2 图像坐标系统示例

$I(r,c)$ 是对 $f(x,y)$ 的离散化后的结果。 r 表示图像的行 (row); c 表示图像的列 (column); I 表示离散后的 f ; I,r,c 的值只能是整数。数字图像可用矩阵或数组进行描述。

图像处理可以分为模拟图像处理和数字图像处理。模拟图像处理是指实时光学(实时,速度快)但只是有限处理。例如:望远镜、显微镜、哈哈镜、透镜、胶片合成照相、凸透镜等这种光学变换相当于对图像的一种实时 FFT 变换。图像处理技术是随着计算机的处理速度和数据量增加发展起来的。数字图像处理就是利用计算机系统对数字图像进行各种目的的处理。由于其使用数字信号处理技术,数字图像处理可对图像进行“精细化”处理,达到像素级别的加工与再现因而具有模拟信号所不能比拟的优势,因此本书的研究对象是数字图像,以后章节中的图像专指数字图像。

1.1.1 图像模型

一幅图像可描述为如下模型:在数字图像领域,将图像看成是许多大小相同、形状一致的像素 (picture element, 简称 pixe) 组成,如图 1.3 所示。

一幅图像可以用二维矩阵加以表示,如图 1.4 所示。数字矩阵中 f 代表该像素彩色或灰度值,脚码代表像素的坐标位置,与矩阵的定义相类似,其中第一个脚标代表行(row) 值,第二个脚标代表列(column) 值。

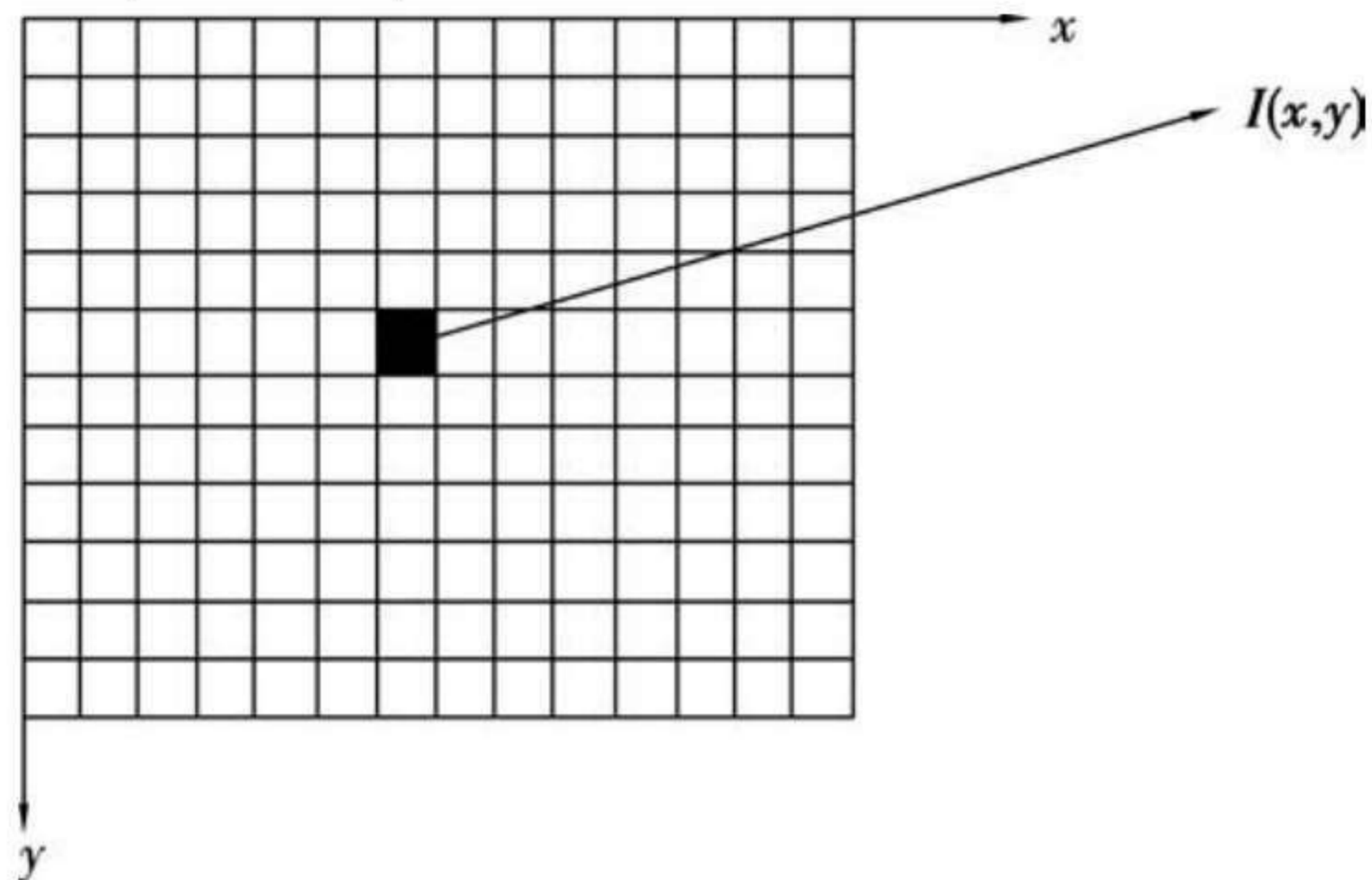


图 1.3 图像数学模型

$$\begin{bmatrix} f_{11} & f_{12} & \dots & f_{1N} \\ f_{21} & f_{22} & \dots & f_{2N} \\ \vdots & \vdots & \dots & \vdots \\ f_{M1} & f_{M2} & \dots & f_{MN} \end{bmatrix}$$

图 1.4 数字图像的数字矩阵表示

举一个简单的例子来说,如图 1.5 所示的灰色图像可以右边的矩阵的值来表示。也就是图像中第一行第 1 像素的灰度值为零,第 2 像素的灰度值为 150,第 3 像素的灰度值为 200,……以此类推,可得到右边的矩阵。

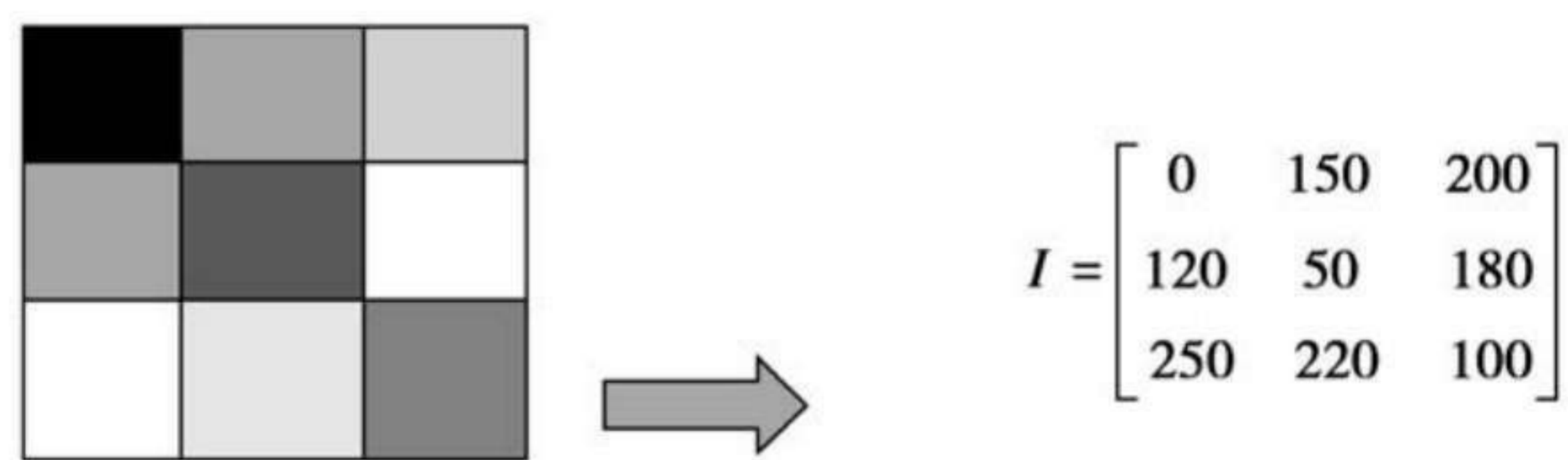


图 1.5 简单的灰色图像及其矩阵表示

但是当图像变得越大,这个矩阵就越复杂。如图 1.6 所示的一幅灰色图像,其内部的一小部分的图像的矩阵就是一个 8×8 的矩阵,如果对整个图像来说,矩阵的数量就比较庞大了,所以数字图像处理中有大量的数据要进行计算和保存,对计算机的要求较高。

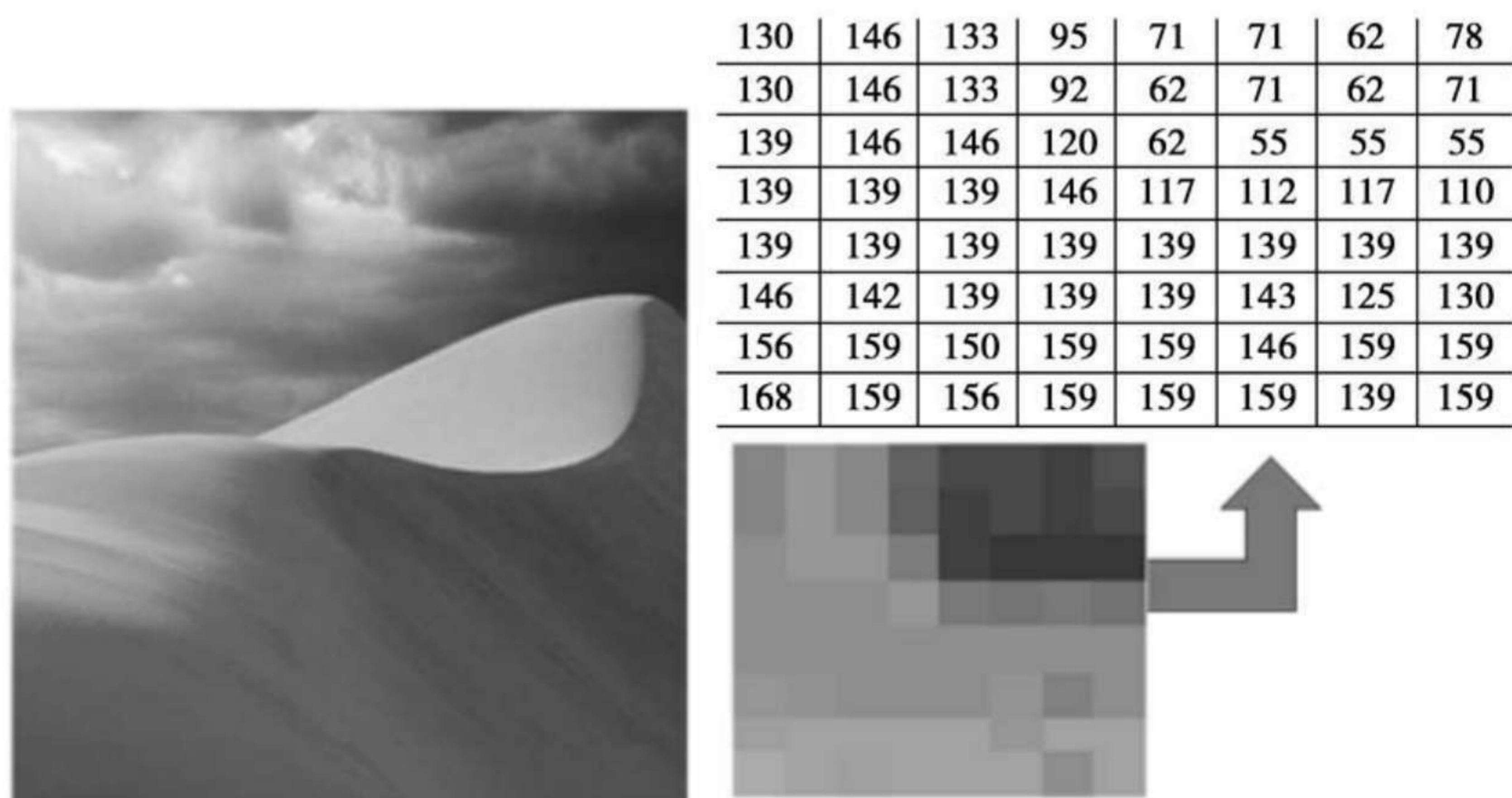


图 1.6 复杂的灰色图像及其矩阵表示

1.1.2 图像工程

图像工程综合了各种图像技术,是集图像研究和应用为一体的整体框架,其研究内容丰富、覆盖范围广,需要利用计算机和其他电子设备完成一系列任务。例如:

- ①提高图像的视觉质量以提供人眼主观满意或较满意的效果。

- ②提取图像中目标的某些特征,以便于计算机分析或机器人识别。
- ③为了存储和传输庞大的图像和视频信息,常常对这类数据进行有效的压缩。
- ④信息的可视化。
- ⑤信息安全的需要。

因此,根据研究内容和研究方法的不同,图像工程可以分为三个层次:

(1) 图像处理

对图像进行各种加工,以改善图像的视觉效果,强调图像之间进行的变换,图像处理是一个从图像到图像的过程。广义上图像处理泛指各种图像技术,但狭义上专指图像的底层基本操作。

(2) 图像分析

对图像中感兴趣的目标进行提取和分割,获得目标的客观信息(特点或性质),建立对图像的描述;以观察者为中心研究客观世界;图像分析是一个从图像到数据的过程。

(3) 图像理解

研究图像中各目标的性质和它们之间的相互联系;得出对图像内容含义的理解及原来客观场景的解释;以客观世界为中心,借助知识、经验来推理、认识客观世界,属于高层操作(符号运算)。

图像处理是相对较低层次的操作,它主要在图像像素级上进行处理,处理的数据量非常大,其重点是图像的变换及其之间的相互转换。在保持视觉效果的基础上减少数据量,为图像的进一步分析和理解奠定基础。图像特征提取是通过各种技术和方法提取图像的视觉特征,把原本以像素描述的图像转变成比较简洁的非图形式的描述,确定图像中目标的特点和固有属性,属于图像处理基础之上的中层操作。图像语义分析是在对图像进行特征提取的前提下,将描述抽象出来的符号进行运算,研究图像中各目标之间的性质及其相互联系,在这个过程中不断引入语义,将计算机对图像底层视觉特征进行处理与分析的结果,用人类思维的语言方式描述出来,得出对图像内容的语义解释。因此,图像语义分析抽象度高,数据量小,属于高层操作。

综上所述,图像特征提取与语义分析是在图像处理的基础上实现的图像较高层次的操作阶段,将原始图像数据经过一系列技术手段操作逐步转化为更有组织和更有用的信息,整个过程是一个复杂的信息处理过程,其处理过程和方法与人类的思维推理有许多类似之处。为了使计算机具有人类的视觉功能,实现计算机中的人类智能,人工智能是必须具备的研究手段,具体包括机器学习、模式识别等方法。将图像中的内容视作一种包含语义的数据,通过研究图像特征提取、目标识别、场景描述等问题,将底层特征转换成有价值的知识描述,给出图像中知识的语义和属性等相关表达,形成图像内容的语义解释体系,最终实现计算机对人类视觉思维信息的模拟过程。

1.1.3 语义分析技术

图像语义是一种关于图像低层特征,低层特征到高层语义的映射,以及高层语义的用户理解范畴的全面表达。图像语义的基本单元是文本描述,其优点是可以描述一些高层的抽象的概念,并且比较直观、容易处理。但是文本描述不容易自动获取,存在内在的主观性,基于语义一般定义的文本表示和图文直接匹配,缺乏对概念之间复杂关系的表达能力,难以独立

完成语义描述的任务。图像语义分析的关键环节是考虑文本之间的相关语义概念,构建语义关联模型,以获得一定的语义分析能力。本节主要介绍结构化以及关联性的语义分析建模技术。

1) 结构化模型

结构化的知识表达模型,主要采用人工智能中诸如语义网络、基于 Agent表示、框架网等方式。这种方式能够表达较为复杂的关系,且具有复杂匹配能力。目前已经有诸多基于下文相关模型对多媒体信息进行组织和管理的成果应用,包括:统一资源描述框架 RDF(Resource Description Framework)、网络本体语言 OWL(Web Ontology Language)以及 MPEG-7 在多媒体内容的元数据以及服务描述的定义等。这些方法在不同的场合具有不同的语义表达与分析的能力,但是目前还没有在不同情况下使用性都很好的通用方法。

2) 基于学习的建模

(1) 聚类分析

聚类分析是通过分析研究对象的特征,将研究对象进行分类的处理过程,涉及多元分析技术和算法。聚类一般分为以下四个步骤:首先,根据研究目的,选择合适的聚类变量;其次,根据相似度量准则进行相似度计算;然后,选定聚类方法进行聚类;最后,评估聚类结果。

关于图像聚类,学术界已经展开了大量相关研究,取得了一定研究成果。目前有很多有效的聚类技术,如层次聚类算法、混合溶解算法、模式搜寻算法、基于分割的算法、最近相邻聚类、模糊聚类和进化聚类方法等。图像聚类通过有监督和无监督的学习将图像划分到若干个语义类,聚类过程中最重要的问题是选择最合适的语义类的数量。Demimen提出了一种选择合适分类数的准则,根据树状结构图来分类,要求分类数既不能过多也不能过少,并且各类重心之间距离必须尽可能大,使得任何类在邻近各类中是最突出的。确定分类数的问题已成为各种聚类方法的主要问题之一,也是聚类分析中迄今为止尚未完全解决的问题之一,主要的障碍是难以对类的结构和内容给出一个统一的定义。实际应用中人们主要根据研究的目的,从实用角度出发,选择合适的分类数。层次聚类法(Hierarchical Cluster)基于分析得到的聚集进程表(agglomeration Schedule)和树状图,来辅助确定分类数,但其缺点是内存和时间开销大,特别是对海量数据进行聚类分析时。迭代聚类法克服了这两个缺点,具有占用内存小、速度快的优点,但缺点是对初始分类非常敏感,通常只能得到局部最优解。现在聚类分析的趋势是把这两种方法结合起来使用,取长补短。首先使用层次聚类法确定分类数,去除奇异值,并重分类得到各类重心,作为迭代聚类法的初始分类中心,这样克服了层次聚类法只能单方向进行聚类的缺点。

(2) 本体论

本体论是关于词汇或概念的理论,给出构成相关领域词汇的基本术语和关系,以及利用这些术语和关系构成的确定词汇外延的有关规则的定义。其目标是捕获相关的领域知识,提供对该领域知识的共同理解,确定领域内通用的词汇,并给出这些词汇(术语)和词汇之间相互关系的明确定义。一般来说,本体提供一组术语和概念来描述某个领域,知识库则使用这些术语来表达该领域的事实。语义分析通常使用本体描述语言来描述语义实体、属性和关系,本体描述语义的推理能力为在图像分析中进行推理分析提供了条件。一类非常典型的推理应用就是在本体知识库已经建立的情况下,用户可以使用类和属性的继承关系对查询进行扩展,提高分析效率。例如,分析与“汽车”相关的文档,在知识库中,由于“越野车”是“汽车”

的子类,所以,可将“越野车”在后台作为扩展的关联术语,扩大分析范围。

目前已有大量关于本体论的图像语义分析的学术研究,但这些研究多是针对特定领域,并且普遍存在创建、更新、维护代价高,可扩展性不强等实用弱点。

(3) 词典

词典是一种高度形式化的、通用、跨语言的知识表示方式,包含了词语的意义信息和词语之间的关系,它不是特定于某个领域的。WordNet就是一种基于认知语言学的英语词典,按照单词的意思组成一个“单词的网络”。它是以同义词集合(synset)作为基本建构单元进行组织的,每个同义词集都代表一个基本的语义概念,并且同义词集合之间以一定数量的关系类型相关联的,一个多义词将出现在它每个意思的同义词集合中。名词、动词、形容词和副词各自被组织成一个同义词的网络。

文本是对图像进行解释的最简单的图像语义表示方法,利用WordNet将文本表示的相关语义概念联系起来,可以获得一定的模糊图文匹配能力。但是WordNet概念太详细和专业,很多概念并不具备普适性,而且缺乏有关两个相关词之间语义距离的信息,不能准确反映词语之间的相似距离。尽管WordNet提供了详细的概念和及其关系描述,却不足以支持关系推理,对于一些概念之间的复杂关系缺乏足够的表达能力,不具有普遍意义。

(4) 图模型

图模型可以方便地对数据样本(即图像像素)及其关联的语义关系进行建模,因而被大量应用在基于统计和机器学习的语义分析方法中。当图的结点代表的图像像素均无标记(label)时,图模型可用来进行无监督聚类。图聚类最经典的方法是图切割(graph cut)算法,其目标是得到图的最优划分(partition),每个划分具有各自的标记类别。基于谱图理论,将求解最优划分问题转化为求解带约束的二次型优化问题,并最终转化为图拉普拉斯(Graph Laplacian)特征向量求解问题,其经典算法为谱聚类(spectral clustering)算法。一般的图模型假设图上只包含单一类型结点(对应单一类型数据样本),而当图上存在两种不同类型结点时可用二部图对不同结点之间的异构关联进行建模。Dhillon提出了一种基于二部图(bipartite graph)的协同谱聚类(spectral co-clustering)方法,并应用到文档聚类中,使得可同时对文档和单词进行聚类。

基于图的多视图学习利用多个图模型对数据实例的多种表示及其关联关系进行表达,并通过图拉普拉斯融合或在图上定义混合Markov链进行学习,相关算法已经应用到了图像的语义分析领域。在图像的异构特征表达中,同一张图像可以提取多种异构视觉特征,因此可以有多种特征表示,如颜色、纹理和形状。在多模态融合中,多模态数据往往有不同表示或来自不同数据空间,如视频镜头可用关键帧视觉特征、转录文本特征、声音特征等多模态特征进行表示。多视图学习(multi view learning)假设每个数据样本有多种表示,这些表示可属不同的空间或数据域。不同的数据表示有其数据域或空间的特定统计特性,并且这些特性之间关联互补。多视图学习在考虑每个视图统计特性的基础上学习得到一个最优的共同表达,以充分利用多种表示的互补特性。

1.2 图像采集与处理

对图像的特征提取与语义分析是以数字图像为研究对象,依靠计算机来观察和认知世界。图像的采集与处理的技术和方法,是反映场景内容的前提条件。本节主要介绍图像的获取技术、存储方式以及图像处理所需的硬件与软件,并着重介绍了实际应用中常用的Matlab操作。

1.2.1 图像的获取技术

1) 图像采集系统

图1.7是图像采集系统原理框图,它可以分成照明系统、同步系统、扫描系统、光/电转换系统、A/D转换系统5部分。

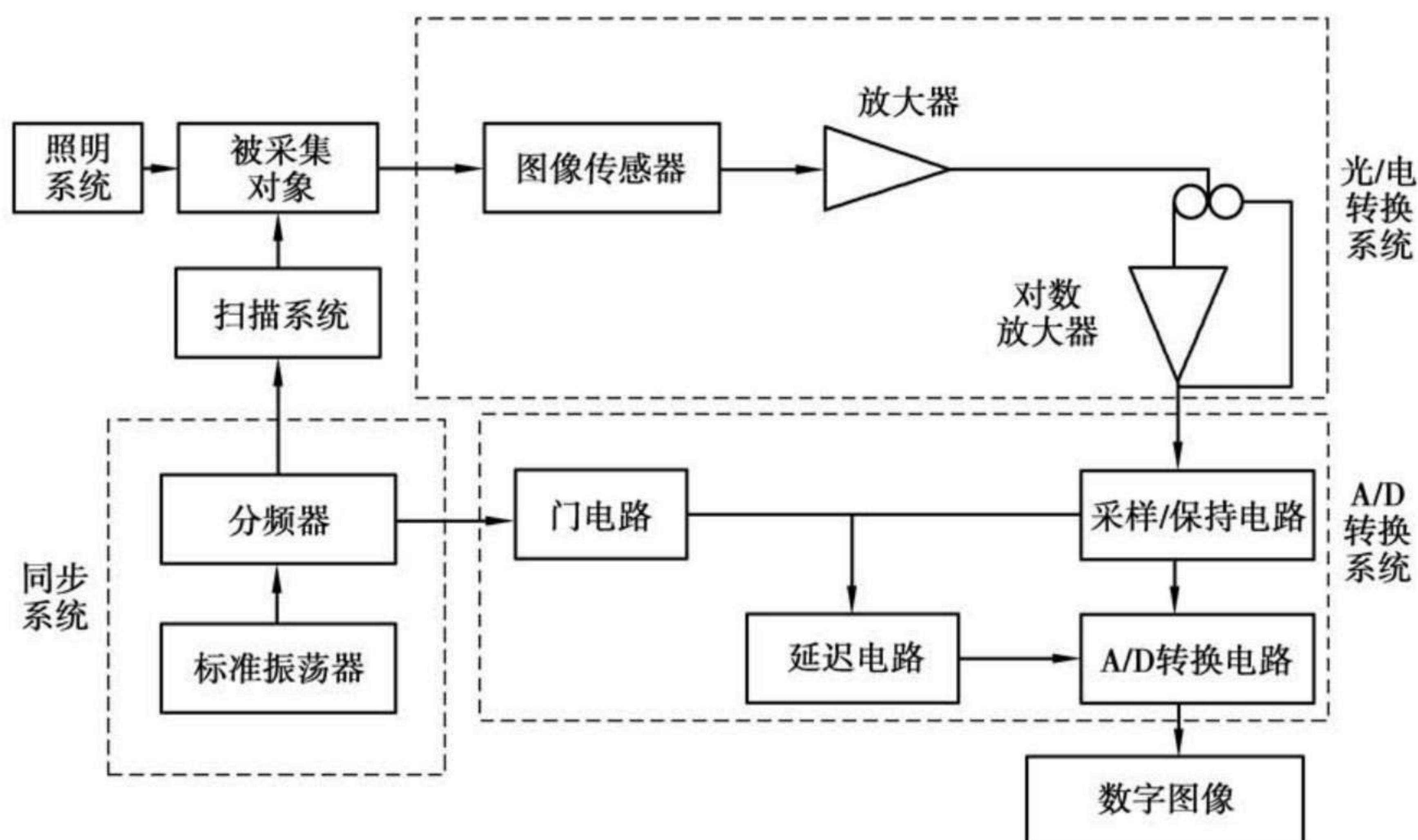


图 1.7 图像采集系统原理框图

照明系统提供光源照射被采集对象(景物),为光/电转换系统提供足够亮度的光强度信号。同步系统提供整个图像采集系统的时钟同步信号,以使系统中的所有部件同步动作。扫描系统是图像采集系统的固有部分,它通过对整幅图像的扫描实现被采样图像空间坐标的离散化,并获得每一个采样点的光强度值。扫描可以采用机械手段、电子束或者集成电路来完成。光/电转换系统负责把扫描系统输出的与采样点属性对应的光信号转换为电信号,并提供必要的放大处理,以与A/D转换系统相匹配。从光/电转换系统输出的电信号进入A/D转换系统,经过采样/保持,A/D转换后,转换成数字信号输出,供存储、显示、传输和其他处理。

(1) 光/电转换特性

图像传感器通过光/电器件将光信号转换为电信号。在照明系统的照射下,如果光信号的能量(光强度)低于光/电器件的感应阈值,光/电器件对该强度的光信号没有反应,称为无感应区域;当光强度达到一定的强度以后,再增加输入的光信号强度,光/电器件产生的电信号强度也不会变化,称为饱和区域;介于无感应区域和饱和区域之间的光强度区域,称为动态

区域。光电器件应该正常工作在动态区域。图 1.8 显示了光/电器件的输入/输出变换特性曲线。

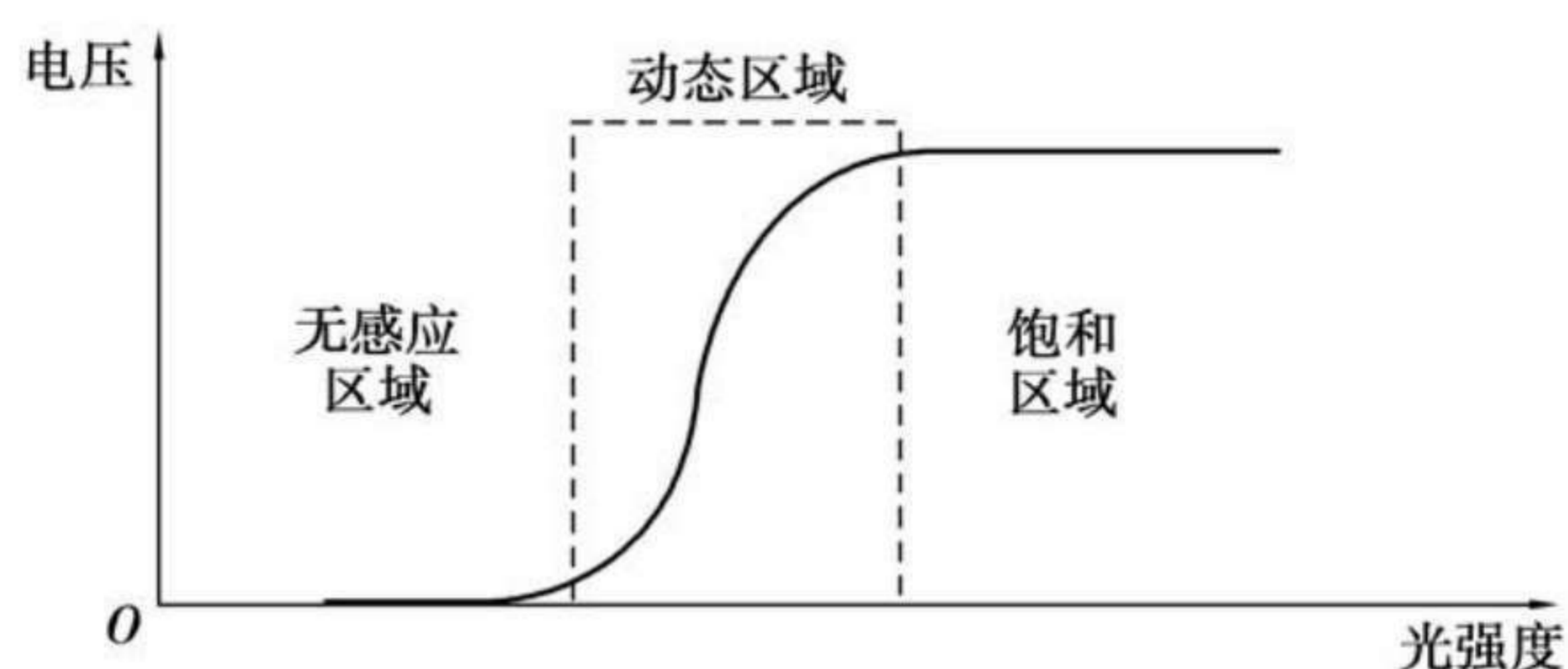


图 1.8 输入/输出变换特性曲线

彩色图像输入时,需要先用分光镜、滤色片等装置对彩色信号进行分解,得到红、绿、蓝三色通道,然后分别对这三个颜色通道进行光/电转换和 A/D 转换。

(2) 图像传感器

图像传感器,图像传感器主要完成光/电转换功能。根据元件的不同,可分为 CCD (Charge Coupled Device, 电荷耦合元件) 和 CMOS (Complementary Metal Oxide Semiconductor, 金属氧化物半导体元件) 两大类。

CCD (Charged Coupled Device) 于 1969 年在贝尔实验室研制成功,之后由日商等公司开始量产,其发展历程已经 40 多年,从初期的 10 多万像素已经发展至目前主流应用的 500 万像素。CCD 又可分为线阵 (Linear) 与面阵 (Area) 两种,其中线阵应用于影像扫描器及传真机上,而面阵主要应用于数码相机、摄录影机、监视摄影机,但两者的工作原理相同。CCD 器件包含若干个光敏像元,每个像元就是一个光敏二极管。被摄物体的反射光线经光学系统聚焦到 CCD 的像敏面上,像敏面将照射在每一像元上的光信号转换为电荷信号存储在像元单元上,再转移到 CCD 的移位寄存器中,在驱动脉冲的作用下从器件中移出成为电信号。CCD 传感器具有高量子效应、优异的电荷传递性能、高占空因数、低噪声、小像素等优点,使其成为目前图像传感器采用的主要技术。迄今为止,CCD 是主要的实用化的固态图像传感技术,数码相机的图像传感器大多以 CCD 为主。

CMOS 图像传感器于 20 世纪 80 年代发明以来,由于当时 CMOS 工艺制造的技术不高,以致传感器在应用中的杂信号较大,商品化进程一直较慢。时至今日,CMOS 传感器的应用范围也开始非常的广泛,包括数码相机、PC Camera、影像电话、第三代手机、视频会议、智能型保全系统、汽车倒车雷达、玩具,以及工业、医疗等用途。在低档产品方面,其画质质量已接近低档 CCD 的解析度,相关业者希望用 CMOS 器件取代 CCD 的努力正在逐渐明朗。与 CCD 相比,CMOS 具有体积小,耗电量不到 CCD 的 1/10,售价也比 CCD 便宜 1/3 的优点。

目前,CMOS 图像传感器可以在低档和中档成像系统方面与 CCD 相媲美;但在高档成像系统方面,CCD 尚未受到 CMOS 的任何挑战。

2) 数字照相机

数码照相机又称数字照相机,是 20 世纪末开发出的新型照相机。在拍摄和处理图像方有着得天独厚的优势。随着电脑的普及以及对电脑图像处理技术的认同,数码照相机在视检测方面得到了广泛的应用。数字照相机由镜头、感光器件 (CCD 或 CMOS)、MPU (微处理器)、内置存储器、LCD (液晶显示器)、PC 卡 (可移动存储器) 和接口 (计算机接口、电视机接口) 等部分组成,其组成框架如图 1.9 所示。

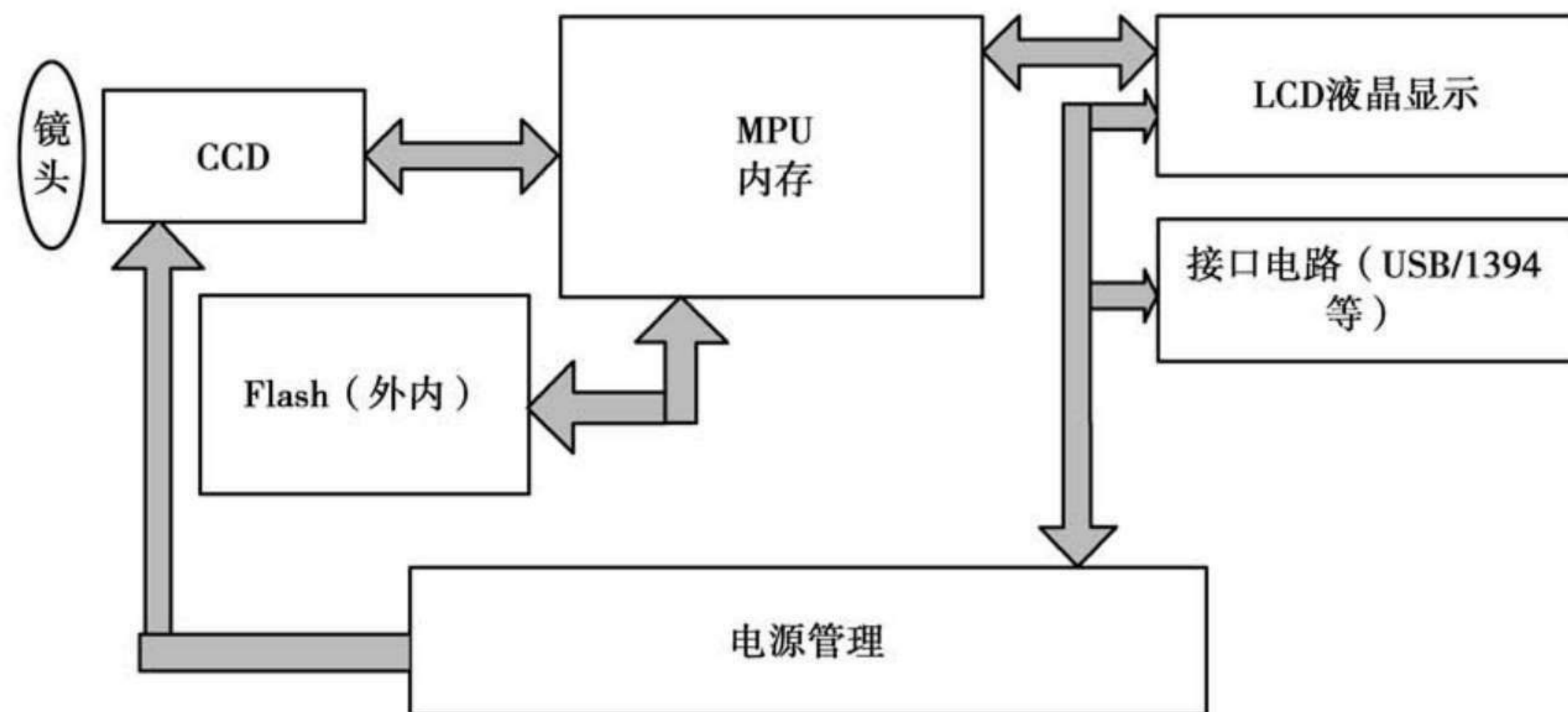


图 1.9 数码照相机组成框图

数码照相机图像存储一般是由系统内置的微处理器来完成。压缩处理与存储图像所用的时间不可忽略,因此在使用数码相机时可以明显感到较长的等待时间。图像格式的种类繁多,一般常用 JPEG 格式。数码照相机的存储器分为内置存储器和可移动存储器。内部存储器为半导体存储器,安装在相机内部,用于临时存储图像,接口传送。可移动存储器 Compact Flash 卡 (CF) 和 SmartMedia 卡等。数码相机输出接口包括计算机通信接口,例如串行接口、并行接口、USB 接口和 SCSI 接口,若使用红外线接口,则要为计算机安装相应的红外接收器及其驱动程序连接电视机的视频接口连接打印机的接口。

3) 数字摄像机

数码摄像机工作的基本原理简单地说就是光电数字信号的转变与传输。即通过光学系统(主要指镜头系统)再通过摄像管或固体摄像器件将光信号转变成电流,再将模拟电信号转变成数字信号,由专用电路进行处理和过滤后得到的信息还原出来就是我们看到的动态画面了。数码摄像机的感光元件能把光线转变成电荷,通过模/数转换器芯片转换成数字信号,主要有两种:一种是广泛使用的 CCD 元件,另一种 CMOS 器件。

4) 图像采集卡

图像采集卡,其功能是将图像信号采集到电脑中,以数据文件的形式保存在硬盘上。图像采集卡是图像采集部分和图像处理部分的接口。图像经过采样、量化以后转换为数字图像并输入、存储到帧存储器的过程,叫作采集。由于图像信号的传输需要很高的传输速度,通用的传输接口不能满足要求,因此需要图像采集卡。图像采集卡还提供数字 I/O 的功能。由于通过高速 PCI 总线可实现直接采集图像到 VGA 显存或主机系统内存,这不仅可以使图像直接采集到 VGA,实现单屏工作方式,而且可以利用 PC 机内存的可扩展性,实现所需数量的序列图像逐帧连续采集,进行序列图像处理分析。此外,由于图像可直接采集到主机内存,图像处理可直接在内存中进行,因此图像处理的速度随 CPU 速度的不断提高而得到提高,因而使得对主机内存的图像进行并行实时处理成为可能。

摄像头实时或准时采集数据,经 A/D 变换后将图像存放在图像存储单元的一个或三个通道中,D/A 变换电路自动将图像显示在监视器上。通过主机发出指令,将某一帧图像静止在存储通道中,即采集或捕获一帧图像,然后可对图像进行处理或存盘。高档图像采集卡还包括卷积滤波、FFT(快速傅里叶变换)等图像处理专用的快速部件。现在有的图像采集卡将图像和图形功能合为一体。这种卡基于 PCI 总线设计,它将图像和 VGA 的图形功能合为一体,可在计算机屏幕上实时显示彩色活动图像。

5) 扫描仪

扫描仪是一种光、机、电一体化的高科技产品,它是将各种形式的图像信息输入计算机的重要工具,通过捕获图像并将之转换成计算机可以显示、编辑、存储和输出的数字化输入设备。对照片、文本页面、图纸、美术图画、照相底片、菲林软片,甚至纺织品、标牌面板、印制板样品等三维对象都可作为扫描对象,提取和将原始的线条、图形、文字、照片、平面实物转换成可以编辑及加入文件中的装置。目前扫描仪的价格并不昂贵,而且种类繁多,但不同的扫描仪将提供不同的图像质量,这正如不同类型的照相机照出不同质量的相片一样。

在开始扫描之前,必须知道自己最终图像的大小,并计算出正确的扫描分辨率。分辨率是扫描仪最主要的技术指标,它表示扫描仪对图像细节的表现能力,即决定扫描仪所记录图像的细致度,其单位为 PPI(Pixels Per Inch)。目前大多数扫描的分辨率为 300 ~ 2 400 PPI。一个图像所包含的像素越多,表明它所容纳的信息也就越多。因此,通常往一个图像填塞的像素越多,图像也就越清晰。PPI数值越大,扫描的分辨率越高,扫描图像的品质越高,但这是有限度的。当分辨率大于某一特定值时,只会使图像文件增大而不易处理,并不能对图像质量产生显著的改善。如果以低分辨率进行扫描,则图像就可能会模糊不清,或者可能会看见图像中单个的像素元素。扫描分辨率一般有两种:真实分辨率(又称光学分辨率)和插值分辨率。光学分辨率就是扫描仪的实际分辨率,它决定了图像的清晰度和锐利度的关键性能指标。插值分辨率则是通过软件运算的方式来提高分辨率的数值,即用插值的方法将采样点周围遗失的信息填充进去,因此也被称作软件增强的分辨率。例如扫描仪的光学分辨率为 300 PPI,则可以通过软件插值运算法将图像提高到 600 PPI,插值分辨率所获得的细部资料要少些。尽管插值分辨率不如真实分辨率,但它却能大大降低扫描仪的价格,且对一些特定的工作例如扫描黑白图像或放大较小的原稿时十分有用。图像的文件大小与图像的分辨率直接相关。一幅以高一些的分辨率扫描的图像所产生的文件比低一些分辨率扫描的图像的文件要大。如果拿来一幅 72 DPI(Dot Per Inch,是打印分辨率使用的单位,意思是:每英寸所表达的打印点数)的图像,然后以 2 倍于原来分辨率大小的分辨率(144 DPI重新扫描,则所得到的新文件就大约是初始文件的 4 倍大小。这样,在扫描时,如果使用的分辨率太高,则图像的文件大小就可能会超过计算机内存的容量。

灰度级表示图像的亮度层次范围。级数越多扫描仪图像亮度范围越大、层次越丰富,目前多数扫描仪的灰度为 256 级。256 级灰阶中以真实呈现出比肉眼所能辨识出来的层次还多的灰阶层次。

色彩数表示彩色扫描仪所能产生颜色的范围。通常用表示每个像素点颜色的数据位数即比特位(bit)表示。所谓 bit这是计算机最小的存储单位,以 0 或 1 来表示比特位的值,越多的比特位数可以表现越复杂的图像信息。例如,常说的真彩色图像指的是每个像素点由三个 8 比特位的彩色通道所组成即 24 位二进制数表示,红绿蓝通道结合可以产生 $2^{24} = 16.67 \text{ M}$ (兆)种颜色的组合,色彩数越多扫描图像越鲜艳真实。

扫描仪可分为三大类型:滚筒式扫描仪和平板扫描仪,近几年才有的笔式扫描仪、便携式扫描仪、馈纸式扫描仪、胶片扫描仪、底片扫描仪和名片扫描仪。

1.2.2 图像存储方式

1) 图像存储设备

数字图像处理用于处理短期快速存储,常用的是计算机内存,另一种是帧存。图像帧存储体(frame buffer或 frame store memory),简称为帧存(FB),是指至少能存储一帧电视图像,并能按照一定的速率显示所存储图像的由半导体存储器芯片组成的存储装置。计算机显存和图像卡上的存储数字图像的半导体缓存,都可以称为帧存。帧存主要用于解决快速的视频数据流(如 8 bit/100 ns)和计算机慢速的数据传输速率之间的矛盾。考虑到存储的速度,快速调用的在线存储,常采用磁盘,现在也有采用固态硬盘作为存储介质。

2) 常见的图像存储格式

数字图像有多种存储格式,在计算机中是以图像文件的形式存放的,每种格式一般由不同的开发商支持。随着信息技术的发展和图像应用领域的不断拓宽,还会出现新的图像格式。因此,要进行图像处理,必须了解图像文件的格式,即图像文件的数据构成。每一种图像文件均有一个文件头,在文件头之后才是图像数据。文件头的内容由制作该图像文件的公司决定,一般包括文件类型、文件制作者、制作时间、版本号、文件大小等内容。目前较常用的静态图像文件格式有 BMP、GIF、TIFF、JPEG 等类型。

(1) BMP格式

BMP 又称位图文件(bitmap, BMP),是一种与硬件设备无关的图像文件格式,使用非常广。它采用位映射存储格式,除了图像深度可选以外,不采用其他任何压缩,因此, BMP 文件所占用的空间很大。BMP 文件的图像深度可选 1 bit、4 bit、8 bit 及 24 bit。BMP 文件存储数据时,图像的扫描方式是按从左到右、从下到上的顺序。它是 Windows 软件推荐使用的一种格式,随着 Windows 的普及, BMP 文件格式的应用越来越广泛。BMP 格式的文件名后缀是“. bmp”,它的色彩深度有 1 位、4 位、8 位及 24 位几种格式。

由于 BMP 文件格式是 Windows 环境中交换与图有关的数据的一种标准,因此在 Windows 环境中运行的图形图像软件都支持 BMP 图像格式。典型的 BMP 图像文件由三部分组成:位图文件头数据结构,它包含 BMP 图像文件的类型、显示内容等信息;位图信息数据结构,它包含有 BMP 图像的宽、高、压缩方法以及定义颜色等信息,位图文件参数头的含义如表 1.1 所示。

表 1.1 位图文件参数头

字节数	参 数	说 明
2	bftype	文件类型,以 BM 标识
4	bfsz	实际图像数据长度
2	reserved1	保留
2	reserved1	保留
4	offset	文件开始到位图数据开始处的偏移量

位图参数头域含有关于这幅图像的信息,例如:以像素为单位的宽度和高度,以及位图的彩色、压缩方法等,其结构如表 1.2 所示。