

1

绪论

本章先对全书内容要点和布局结构进行概括介绍,将如下安排:

1.1 节首先介绍图像和数字图像的基本概念,以及对图像的表达(包括图像的表达方法和图像的显示形式)。

1.2 节将全面介绍对整个图像领域进行综合集成的研究和应用的新学科——图像工程。图像工程的提出是图像技术发展的必然,目前包括三个层次,与多个学科有密切关系,也有广泛的应用。为对图像工程研究应用的发展和现状有一个全面的了解,还介绍了对图像工程文献统计分类的一些结果。

1.3 节结合对图像处理系统框架的介绍,讨论完成其中各个模块的功能的设备情况。本书主要涉及图像处理的原理和算法,但硬设备是构成系统和实现各种算法的基础。

1.4 节列出本书主要内容、结构和安排,并对各章重要内容给予概括介绍。

1.1 图像基本概念

1.1.1 图像和数字图像

1. 图像

人们对图像是很熟悉的,生活中人们很容易说出哪些东西是图像。一般将图像看作对物体或场景的一种表现形式,是用来表示某种东西的一种东西。例如,字典里对图像的一个定义是“对物体的表达、表象、模仿,一个生动的视觉描述,为了表达其他事物而引入的事物”[Bow 2002]。严格一点说,图像是用各种观测系统以不同形式和手段观测客观世界而获得的,可以直接或间接作用于人眼并进而产生视知觉的实体[章 1996a]。人的视觉系统就是一个观测系统,通过它得到

的图像就是客观景物在人心目中形成的影像。

图像包含了它所表达的物体的描述信息。我们生活在一个信息时代,科学研究和统计表明,人类从外界获得的信息约有 75% 来自视觉系统,也就是从图像中获得的。这里图像的概念是比较广义的,包括照片、绘图、动画、视像,甚至文档等等。中国有句古话,“百闻不如一见”。人们常说,“一图值千字”。它们都说明图像中所含的信息内容非常丰富,而事实上图像也确实带有大量的信息,是我们最主要的信息源。

本书主要讨论对自然场景成像得到的图像,这也有许多类别。例如用数码相机拍摄的照片(人,风景等),用数码摄像机拍摄的视频(家庭联欢,足球比赛等),用监控系统记录下的各种序列(交通管理,导弹飞行等),用太空望远镜摄取的各种电磁辐射图像,用雷达依靠反射波形成的图像,医学上常使用的 X 光图像、B 超图像、CT 图像,磁共振图像等。图像不仅有灰度的和彩色的,还可以有纹理的,深度的等。

客观世界在空间上是三维(3-D)的,但一般从客观景物得到的图像是二维(2-D)的。一幅图像可以用一个 2-D 数组 $f(x, y)$ 来表示,这里 x 和 y 表示 2-D 空间 XY 中一个坐标点的位置(实际图像的尺寸是有限的,所以 x 和 y 的取值也是有限的),而 f 则代表图像在点 (x, y) 的某种性质 F 的数值(实际图像中各个位置上所具有的性质 F 的取值也是有限的,所以 f 的取值也是有限的)。例如常用的图像一般是灰度图,这时 f 表示灰度值,它常对应客观景物被观察到的亮度。文本图像常为二值图像, f 的取值只有两个。图像在点 (x, y) 也可有多种性质,此时可用矢量 $\mathbf{f}$ 来表示。例如一幅彩色图像在每一个图像点同时具有红绿蓝 3 个值,可记为 $[f_r(x, y), f_g(x, y), f_b(x, y)]$ 。需要指出,我们总是根据图像内不同位置的不同性质来利用图像的。

例 1.1.1 一般的图像表达函数

图像可表示一种辐射能量的空间分布,这种分布可以是 5 个变量的函数, $T(x, y, z, t, \lambda)$, 其中 x, y, z 是空间变量, t 代表时间变量, λ 是频谱变量(波长)。例如,红色的物体反射波长在 $0.57\mu\text{m}$ 到 $0.70\mu\text{m}$ 的光而且吸收几乎所有其他波长的能量,绿色的物体反射波长在 $0.48\mu\text{m}$ 到 $0.57\mu\text{m}$ 的光,蓝色的物体反射波长在 $0.40\mu\text{m}$ 到 $0.48\mu\text{m}$ 的光,紫外(色)的物体反射波长在 $0.25\mu\text{m}$ 到 $0.40\mu\text{m}$ 的光,红外(色)的物体反射波长在 $0.9\mu\text{m}$ 到 $1.5\mu\text{m}$ 的光。它们合起来覆盖了波长为 $0.25\mu\text{m} < \lambda < 1.5\mu\text{m}$ 的范围。由于实际图像在时空上都是有限的,所以 $T(x, y, z, t, \lambda)$ 是一个 5-D 有限函数。□

2. 数字图像

日常所见的图像有许多是连续的,即 f, x, y 的值可以是任意实数。为了能用计算机对图像进行加工,需要把连续的图像在坐标空间 XY 和性质空间 F 都离散化。这种离散化了的图像是数字图像,是客观事物的可视数字化表达。数字图像可以用 $I(r, c)$ 来表

示,其中 I, c, r 的值都是整数。这里 I 代表离散化后的 f , (r, c) 代表离散化后的 (x, y) , 其中 r 代表图像的行(row), c 代表图像的列(column)。本书以后主要讨论数字图像,在不至引起混淆的情况下用 $f(x, y)$ 代表数字图像,如不作特别说明, f, x, y 都在整数集合中取值。

早期英文书籍里一般用 picture 来指图像,随着数字技术的发展,现都用 image 代表离散化了的数字图像,因为“计算机存储人像或场景的数字图像(computers store numerical images of a picture or scene)”[Zhang 1996]。这样看来,应该使用“数字图像”而不是“数字图象”。事实上图象比图像的含义更广,覆盖面更宽。图像中每个基本单元叫做图像元素,在早期用 picture 表示图像时就称为像素(pixel)。对 2-D 图像,英文里常用简称 pixel 代表像素。如果采集一系列的 2-D 图像或利用一些特殊设备还可得到 3-D 图像。对 3-D 图像,英文里常用 voxel 代表其基本单元,简称体素(volume element)。近年由于都用 image 代表图像,所以也有人建议用 imel 统一代表像素和体素。

1.1.2 图像的表达式

1. 图像的矩阵和矢量表示

前面提到,一幅 2-D 图像可以用一个 2-D 数组 $f(x, y)$ 来表示。实际中还常将一幅 2-D 图像写成一个 2-D 的 $M \times N$ 矩阵(其中 M 和 N 分别为图像的总行数和总列数):

$$\mathbf{F} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1N} \\ f_{21} & f_{22} & \cdots & f_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ f_{M1} & f_{M2} & \cdots & f_{MN} \end{bmatrix} \quad (1.1.1)$$

一幅 2-D 图像也可以用矢量来表示,可写成:

$$\mathbf{F} = [f_1 \quad f_2 \quad \cdots \quad f_N] \quad (1.1.2)$$

其中:

$$f_i = [f_{1i} \quad f_{2i} \quad \cdots \quad f_{Mi}]^T \quad i = 1, 2, \dots, N \quad (1.1.3)$$

上述两种表示形式可以方便地互相转换。

2. 图像的显示方式

对 2-D 图像的显示可以采取多种形式,其基本思路是将 2-D 图像看作在 2-D 空间位置上的一种幅度分布。根据图像的不同,采取的显示方式也可不同。例如对二值图像,在每个空间位置的取值只有两个,可用黑白来区别,也可用 0 和 1 来区别。

图 1.1.1 给出对同一个 2-D 的二值图像矩阵的 3 种不同的可视表达方式。在图像表达的数学模型中,一个像素区域常用其中心来表示,这样得到的表达形式就是平面上的离散点集,对应图 1.1.1(a)。如果将像素区域仍用区域来表示,这就得到如图 1.1.1(b)的图形。当把幅度值标在图像中相应的位置,就得到如图 1.1.1(c)的类似矩阵的结果。用

图 1.1.1(b) 的形式也可表示有多个灰度的图像,此时需要用不同深浅的色调表示不同的灰度。用图 1.1.1(c) 的形式也可表示有多个灰度的图像,此时将不同灰度用不同的数值表示。

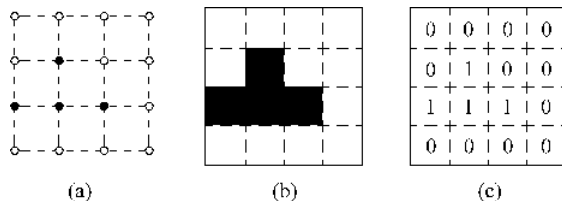


图 1.1.1 3 种表达同一个 4×4 的二值图像矩阵的可视方式

在图 1.1.1 中,各种方式表示的都是一个 4×4 的二值图像,一般说这幅图像的空间分辨率是 4×4 ,也就是说这幅图像在空间有 $4 \times 4 = 16$ 个位置可显示独立的灰度。在显示设备允许的情况下,可对每个空间位置赋以不同灰度来表示图像,见例 1.1.2。这也是最常用的图像表示方式。在显示设备只可直接显示二值时,则可借助 1.3.3 小节的技术来显示灰度图像。对彩色图像的显示,可参见第 7 章的介绍。

例 1.1.2 灰度数字图像显示示例

图 1.1.2 给出 2 幅典型的灰度数字图像。图(a)所用的坐标系常在屏幕显示中采用(屏幕扫描是从左向右,从上向下进行的),它的原点(origin) O 在图像的左上角,纵轴标记图像的列,横轴标记图像的列。 $I(r,c)$ 既可代表这幅图像,也可表示在 (r,c) 行列交点处的图像值。图(b)所用的坐标系常在图像计算中采用,它的原点在图像的左下角,横轴为 X 轴,纵轴为 Y 轴(与常用的笛卡儿坐标系相同)。 $f(x,y)$ 既可代表这幅图像,也可表示在 (x,y) 坐标处像素的值。

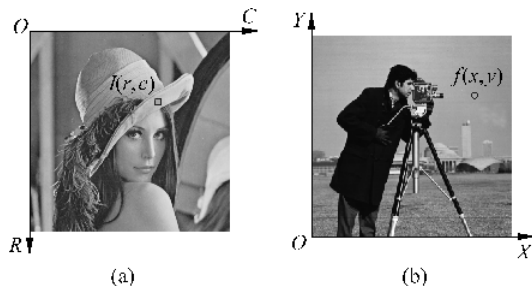


图 1.1.2 数字图像表达示例

□

对灰度图像,为表示不同的灰度级别,可采用 3-D 透视的方法,对图 1.1.2(a) 的两种显示方式见图 1.1.3。图 1.1.3(a) 是将相邻像素的灰度结合起来转换为高程(越浅越高)得到的。这种方法一方面对复杂图像不易表达清楚,另一方面由于透视发生变形看起来

也不够直观。图 1.1.3(b)是将各个像素的灰度显示在对应的高度并保持原灰度得到的,虽较复杂但效果较好。

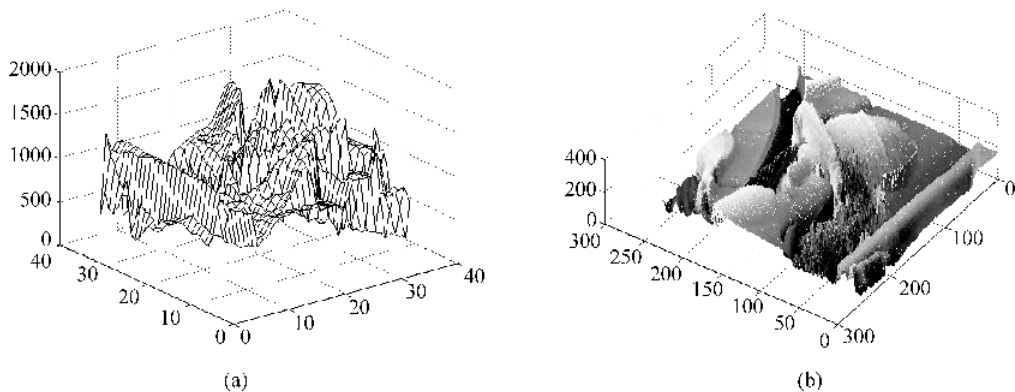


图 1.1.3 数字图像的透视表达示例

1.2 图像工程简介

随着计算机技术的发展,数字图像的采集和应用,特别是加工技术近年来得到极大的重视和长足的进展,出现了许多有关的新理论、新方法、新算法、新手段和新设备,并已使得数字图像技术在科学研究、工业生产、医疗卫生、教育、娱乐、管理和通信等方面得到了广泛的应用,对推动社会发展、改善人们生活水平都起到了重要的作用。

为了对各种图像技术进行综合研究、集成应用,有必要建立一个整体框架——图像工程。图像工程是一门系统地研究各种图像理论、技术和应用的新的交叉学科。它的研究方法与数学、物理学、生理学、心理学、电子学、计算机科学等学科相互借鉴,它的研究范围与模式识别、计算机视觉、计算机图形学等专业互相交叉,它的研究进展与人工智能、神经网络、遗传算法、模糊逻辑等理论和技术密切相关,它的发展应用与生物医学、遥感、通信、文档处理等许多领域紧密结合。

1.2.1 图像技术和图像工程

1. 图像技术及发展历史

图像技术在广义上是各种与图像有关的技术的总称。本套书主要讨论计算机图像技术,这包括利用计算机和其他电子设备进行和完成的一系列工作,例如图像的采集、获取、编码、存储和传输,图像的合成和产生,图像水印的嵌入和提取,图像的显示和输出,图像的变换、增强、恢复(复原)和重建,图像的分割,目标的检测、跟踪、表达和描述,目标特征

的提取和测量,序列图像的校正,3-D 景物的重建复原,图像数据库的建立、索引和抽取,图像的分类、表示和识别,图像模型的建立和匹配,图像和场景的解释和理解,以及基于它们的判断决策和行为规范等等。另外,图像技术还可包括为完成上述功能而进行的硬件设计及制作等方面的技术。随着人们研究的深入和应用的广泛,已有的图像技术在不断更新和扩展,许多新的图像技术也在不断诞生。对各种典型的图像技术的原理和方法的介绍是本书的主要内容。

尽管计算机图像技术的历史可追溯到 1946 年世界上第一台电子计算机的诞生(借助打印设备进行数字图像处理甚至可追溯到 20 世纪 20 年代[Gonzalez 1992]),但在 20 世纪 50 年代计算机主要还是用于数值计算,满足不了处理大数据量图像的要求。在 20 世纪 60 年代,第 3 代计算机的研制成功,以及快速傅里叶变换算法的发现和应用使得对图像的某些计算得以实际实现,人们从而逐步开始利用计算机对图像进行加工利用。在 20 世纪 70 年代,图像技术有了长足的进展,而且第一本重要的图像处理专著[Rosenfeld 1976]也得以出版。在 20 世纪 80 年代,各种硬件的发展使得人们不仅能处理 2-D 图像而且开始处理 3-D 图像。许多能获取 3-D 图像的设备 and 处理分析 3-D 图像的系统研制成功,图像技术得到了广泛的应用。在 20 世纪 90 年代,图像技术已逐步涉及人类生活和社会发展的各个方面。以当时得到广为宣传和应用的多媒体为例,图像在其中其实占据了最主要的地位。广义上来说,文本、图形、视频等都需要借助图像技术才能充分利用。进入 21 世纪,图像技术得到了进一步的发展和应用,在改变人们的生活方式以及社会结构等方面都起到了重要作用。

2. 图像工程的提出

由于图像技术近年来得到极大的重视和长足的进展,出现了许多新理论、新方法、新算法、新手段、新设备、新应用。对各种图像技术进行综合集成的研究和应用应当在一个整体框架下进行,这个框架就是图像工程[章 1996b]。众所周知,工程是指将自然科学的原理应用到工业部门而形成的各学科的总称。图像工程学科则是一个将数学、光学等基础科学的原理,结合在图像应用中积累的经验而发展起来的,将各种图像技术集中结合起来的,对整个图像领域进行研究应用的新学科。事实上,图像技术多年来的发展和积累为图像工程学科的建立打下了坚实的基础,而各类图像应用也对图像工程学科的建立提出了迫切的需要。

1.2.2 图像工程的 3 个层次

图像工程的内容非常丰富,覆盖面也很广,根据抽象程度和研究方法等的不同可分为 3 个层次(见图 1.2.1):图像处理、图像分析和图像理解。换句话说,图像工程是既有联系又有区别的图像处理、图像分析及图像理解 3 者的有机结合,另外还包括对它们的工程应用(见 1.2.4 小节)。

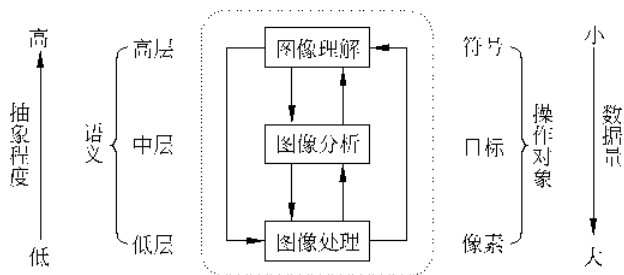


图 1.2.1 图像工程 3 层次示意图

图像处理可看作一大类图像技术,着重强调在图像之间进行的变换。虽然人们常用图像处理泛指各种图像技术,但比较狭义的图像处理技术的主要目标是要对图像进行各种加工以改善图像的视觉效果并为其后的目标自动识别打基础,或对图像进行压缩编码以减少图像存储所需的空间或图像传输所需的时间(从而也降低了对传输通路的要求)。本册书将集中介绍图像处理。

图像分析则主要是对图像中感兴趣的目标进行检测和测量,以获得它们的客观信息从而建立对图像和目标的描述。如果说图像处理是一个从图像到图像的过程,则图像分析是一个从图像到数据的过程。这里数据可以是对目标特征测量的结果,或是基于测量的符号表示。它们描述了图像中目标的特点和性质。本套书中册将集中介绍图像分析。

图像理解的重点是在图像分析的基础上,进一步研究图像中各目标的性质和它们之间的相互联系,并通过对图像内容含义的理解得出对原来客观场景的解释,从而指导和规划行动。如果说图像分析主要是以观察者为中心研究客观世界(主要研究可观察到的事物),那么图像理解在一定程度上是以客观世界为中心,借助知识、经验等来把握整个客观世界(包括没有直接观察到的事物)。本套书下册将集中介绍图像理解。

由上所述,图像处理、图像分析和图像理解分别处在 3 个抽象程度和数据量各有特点的不同层次上,可参见图 1.2.1。图像处理是比较低层的操作,它主要在图像像素级上进行处理,处理的数据量非常大。图像分析则进入了中层,分割和特征提取把原来以像素描述的图像转变成比较简洁的非图形式的描述。图像理解主要是高层操作,基本上是对从描述抽象出来的符号进行运算,其处理过程和方法与人类的思维推理可以有许多类似之处。另外由图 1.2.1 可见,随着抽象程度的提高数据量是逐渐减少的。具体说来,原始图像数据经过一系列的处理过程逐步转化为更有组织和用途的信息。在这个过程中,语义不断引入,操作对象发生变化,数据量得到了压缩。另一方面,高层操作对低层操作有指导作用,能提高低层操作的效能。

1.2.3 图像工程相关学科和领域

图像工程是一门系统地研究各种图像理论、技术和应用的新的交叉学科。从它的研究方法来看,它与数学、物理学、生理学、心理学、电子学、计算机科学等许多学科可以相互借鉴,从它的研究范围来看,它与模式识别、计算机视觉、计算机图形学等多个专业又互相交叉,图 1.2.2 给出图像工程与一些相关学科和领域的联系和区别[章 1996b]。另外,图像工程的研究进展与人工智能、神经网络(一个简单介绍见[章 2000b])、遗传算法、模糊逻辑等理论和技术都有密切的联系,它的发展应用与通信、医学、遥感、文档处理和工业自动化等许多领域也是不可分割的。

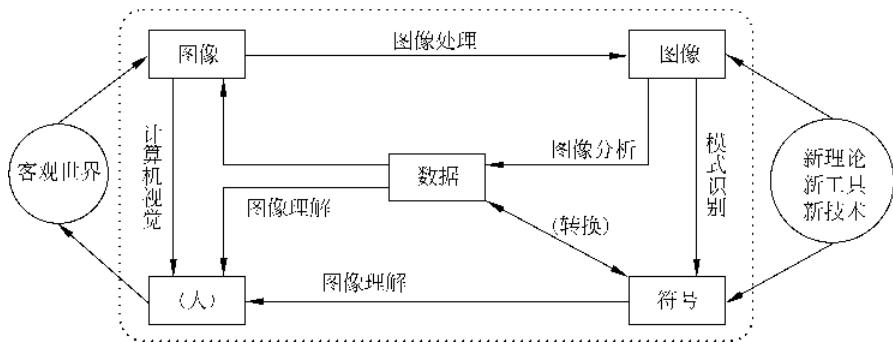


图 1.2.2 图像工程与相关学科和领域的联系和区别

从图 1.2.2 可以看到图像工程 3 个层次各自不同的输入输出内容以及它们与计算机图形学、模式识别、计算机视觉等相近学科的关系。图形学原本指用图形(graph)、图表(chart)、绘图(drawing)等形式表达数据信息的科学,而计算机图形学研究的就是如何利用计算机技术来产生这些形式。如果将它和图像分析对比,两者的处理对象和输出结果正好对调。计算机图形学试图从非图像形式的数据描述来生成(逼真的)图像。另一方面,(图像)模式识别与图像分析则比较相似,只是前者试图把图像分解成可用符号较抽象地描述的种类[章 2000b]。它们有相同的输入,而不同的输出结果可以比较方便地进行转换。至于计算机视觉主要强调用计算机实现人的视觉功能,这中间实际上用到图像工程 3 个层次的许多技术,但目前的研究内容主要与图像理解相结合。

由此看来,以上学科互相联系,覆盖面有所重合。事实上这些名词也常混合使用,它们在概念上或实用中并没有绝对的界限。在许多场合和情况下,它们只是专业和背景不同的人习惯使用的不同术语。它们虽各有侧重但常常是互为补充的。另外以上各学科都得到了包括人工智能、神经网络、遗传算法、模糊逻辑等新理论、新工具、新技术的支持(见图 1.2.2),所以它们又都在近年得到了长足进展。总的来说,图像工程既能较好地许多相近学科兼蓄并容,也进一步强调了图像技术的应用,所以我们选用图像工程来概括整

个图像领域的研究应用。

1.2.4 图像工程的技术应用

近年来数字图像技术已在许多领域得到广泛应用,下面是一些典型的例子:

- (1) 视频通信: 可视电话, 电视会议, 按需电视(VOD), 远程教育;
- (2) 文字档案: 文字识别, 过期档案复原, 邮件分检, 办公自动化, 支票, 签名辨伪;
- (3) 生物医学: 红白学球计数, 染色体分析, X光、CT、MRI、PET 图像分析, 显微医学操作, 对放射图像、显微图像的自动判读理解, 人脑心理和生理的研究, 医学手术模拟规划, 远程医疗;
- (4) 遥感测绘: 矿藏勘探, 资源探测, 气象预报, 自然灾害监测监控;
- (5) 工业生产: 工业检测, 工业探伤, 自动生产流水线监控, 邮政自动化, 移动机器人及各种危险场合工作的机器人, 无损探测, 金相分析, 印刷板质量检验, 精细印刷品缺陷检测;
- (6) 军事公安: 军事侦察, 合成孔径雷达图像分析, 巡航导弹路径规划/地形识别/制导, 无人驾驶飞机飞行, 罪犯脸形合成、识别、查询, 指纹、印章的鉴定识别, 战场环境/场景建模表示;
- (7) 交通: 太空探测, 航天飞行, 公路交通管理, 自动行驶车辆安全操纵。

1.2.5 图像工程文献统计分类

图像工程是一门综合学科, 它的研究内容非常广泛, 覆盖面也很大。从 1996 年起, 《中国图像图形学报》上连续刊登了对图像工程文献统计分类的综述文章[章 1996a]、[章 1996b]、[章 1997a]、[章 1998]、[章 1999a]、[章 2000a]、[章 2001a]、[章 2002a]、[章 2003a]、[章 2004a]、[章 2005]。

表 1.2.1 给出了对从 1995 年到 2004 年期间发表在国内 15 种图像工程重要期刊上的图像工程文献统计分类的概况[章 2005]。表 1.2.1 中, 论文总数表示文献综述系列所选 15 种期刊上发表的论文总数, 选取总数为其中有关图像工程文献的数量。有关文献共分成 5 类, 前 3 类对应图像工程的 3 个层次或 3 大类技术, 第 4 类包括各类技术的技术应用, 第 5 类包括对跨类技术的研究文献。

从表 1.2.1 的统计数据可以看出:

- (1) 这 10 年中无论是论文总数还是选取总数都是逐年增长的。论文总数的增加表明这些刊物自身近年来在不断发展, 而选取总数的增长则表明对图像工程的研究和应用在我国越来越广泛;
- (2) 文献选取率反映了图像工程在各刊覆盖专业范围中的重要性。这 10 年来, 文献选取率基本上是单增的, 图像工程文献在这么多学会和专业的刊物中不断增加, 并已占到

这么大的比例,充分说明图像技术已得到了广泛的重视和应用;

(3) 有关图像处理的文献数量一直都比其他各类的文献数量要多,表明图像处理在图像工程目前发展中仍占有重要地位。

表 1.2.1 1995 年至 2004 年图像工程文献选取统计表

年度	文献总数	选取总数	选取率(%)	图像处理	图像分析	图像理解	技术应用	综述评论
1995	997	147	14.74	35	52	14	46	
1996	1205	212	17.59	52	72	30	55	3
1997	1438	280	19.47	104	76	36	60	4
1998	1477	306	20.72	108	96	28	71	3
1999	2048	388	18.95	132	137	42	73	4
2000	2117	464	21.92	165	122	68	103	6
2001	2297	481	20.94	161	123	78	115	4
2002	2426	545	22.46	178	150	77	135	5
2003	2341	577	24.65	194	153	104	119	7
2004	2473	622	25.60	235	176	76	142	3
小计	18819	4032		1364	1157	553	919	39
平均	1882	403	21.44	136.4	115.7	55.3	91.9	3.9

表 1.2.2 给出了对这些年图像工程有关文献的详细分类统计情况。这里共分了 5 大类 21 小类(其中 A5,B5,C4 是从 2000 年开始增加的,文献数量为近 5 年的统计数据),从中既可对图像工程的内容分类及各种图像技术在我国发展的情况有个概括了解,也可看出当前研究应用的热点和焦点。

表 1.2.2 1995 年至 2004 年图像工程文献分类统计表

大类及名称	文献数量	小类及名称	文献数量
A: 图像处理	1129	A1: 图像采集、获取及存储(包括成像方法、摄像机校正等)	194
		A2: 图像重建(从投影等重建图像)	116
		A3: 图像变换、滤波、增强、恢复/复原、校正等	325
		A4: 图像(视频)压缩编码(包括算法研究、国际标准实现等)	538
		A5: 图像数字水印和图像信息隐藏	191
B: 图像分析	981	B1: 边缘检测、图像分割	477
		B2: 目标表达、描述、测量(包括二值图处理等)	157
		B3: 目标颜色、形状、纹理、空间和运动等的分析	132
		B4: 目标检测、提取、跟踪、识别和分类	259
		B5: 人脸和器官的检测、定位与识别(人体生物特征提取和验证)	132