

# 中国图像工程及当前的几个研究热点

章毓晋

(清华大学电子工程系 北京 100084)

**摘要** 首先对图像工程及其三个层次给予概括介绍,并对中国图像工程文献综述系列的分类结果进行分析,由此归纳出近几年中国图像工程研究和应用的4个研究新热点。这4个处在图像工程三个不同层次的新热点分别是图像数字水印和图像信息隐藏,人脸和器官的检测、定位与识别,多传感器图像信息融合以及基于内容的图像和视频检索。文中对这4个新热点的基本情况、研究现状、技术原理、应用特点、存在问题、发展方向等进行了较为详细的介绍和讨论,为了解图像工程研究应用的情况和开展相关工作提供了参考。

**关键词** 图像工程,文献综述,图像数字水印,人脸识别,图像信息融合,基于内容检索  
**中图法分类号** TN 911.73; TP 391

## Image Engineering in China and Some Current Research Focuses

Zhang Yujin

(Department of Electronic Engineering, Tsinghua University, Beijing 100084)

**Abstract** This paper first gives a compendious introduction of image engineering and its three layers, and then analyses the classification results of the survey series of the yearly bibliographies on image engineering in China. From the analysis, four research hotspots in the research and applications of image engineering in China are induced. These four hotspots, named digital image watermark and information hiding, detection and recognition of human face and organ, multi-sensor data fusion, content-based visual information retrieval, are belong to the three layers of image engineering. The basic principal, current research situation, application particularity, existing problems, development trend and directions are detailed in the paper. This will provide some useful information for the research and application of image engineering and practical references for launching related development works.

**Key words** image engineering, bibliography survey, image watermarking, human face recognition, image data fusion, content-based retrieval

## 1 图像工程

### 1.1 图像工程及三个层次

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,是可以直接或间接作用于人眼并进而产生视觉的实体<sup>[1]</sup>。图像技术在广义上是各种与图像有关的技术的总称。目前人们主要研究的是数字图像,主要应用的是计算机图像技术。

图像技术近年来得到了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。为对它们进行综合研究和集成应用,我们提出了一个整体框架——图像工程<sup>[1]</sup>。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。图像工程学科则是将数学、光

学等基础科学的原理,结合在图像应用中积累的技术经验而发展起来的一个对整个图像领域进行研究应用的新学科。图像工程的内容是非常丰富的,根据抽象程度和研究方法等的不同可分为三个层次:图像处理、图像分析和图像理解,如图1所示。换句话说,图像工程是既有联系又有区别的这三个层次的有机结合,另外还包括对它们的工程应用。

由图1可见,图像处理指比较低层的操作,着重强调在图像之间进行的变换;图像分析则进入了中层,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,从而建立对图像的描述;图像理解主要是高层操作,对从描述抽象出来的符号进行运算,通过研究目标的性质和它们之间的相互联系,得到对图像内容含义的理解以及对原来客观场景的解释,从而指导和规划行动。原始图像数据经过三个层次的加工将逐步转化为更有组织和用途的信息。在这个

过程中,操作对象发生变化,语义不断引入,抽象程度提高,数据量得到了压缩

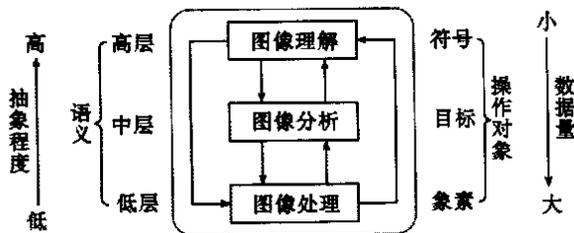


图 1 图像工程三层次示意图

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 2 给出图像工程与相关学科和领域的联系和区别<sup>[2]</sup>。另外,图像工程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都有密切的联系,它的发展应用与医学、遥感、通信、文档处理和工业自动化等许多领域也是不可分割的

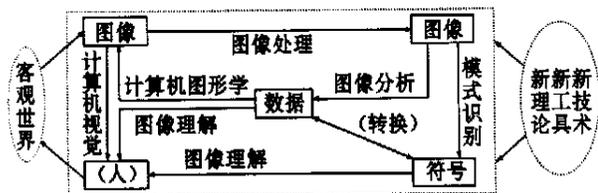


图 2 图像工程与相关学科和领域的联系和区别

从图 2 可以看到图像工程三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

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类别。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三个层次的许多技术,但目前的研究内容主要与图像理解相对应。需要指出,这些学科名词是互相联系的,覆盖面也有所重合,而且近年来有许多相互结合的实例,如基于图像的绘制就是计算机视觉与计算机图形学结合的产物。

### 1.2 中国图像工程文献综述系列

为概括我国图像工程发展情况和研究热点,帮助确定图像工程研究方向和制订工作决策,为从事图像工程研究和图像技术应用的人员查阅有关文献提供方便,并对期刊编者和文献作者提供有用的参考信息,笔者从 1996 年开始,已连续 7 年逐年对前一年发表在国内 15 种重要中文期刊上有关图像工程研究应用的论文(共 2 278 篇),根据其主要内容进行了选取、分类和统计,并对结果进行了分析和讨论,构成一个图像工程的年度文献综述系列<sup>[1-8]</sup>。另外,根据文献计量学的原理和方法<sup>[9]</sup>,该综述系列也对统计分析我国图像工程科技人员的水平、现状和变化等情况提供了有参考价值的数据<sup>[10]</sup>。

表 1 给出综述系列前 5 篇中,即对 1995 年到 1999 年发表的 7 165 篇文献,选出的与图像工程有关的 1 333 篇文献的分类情况,这里共分了 5 大类 18 小类<sup>[6]</sup>。从表 1 可对图像工程的内容分类及上世纪最后 5 年在我国研究应用的情况有个概括了解。

表 1 中国图像工程文献分类情况(1995~ 1999)

大类及名称	文献数量	小类及名称	文献数量	年均数量
A: 图像处理	431	A 1: 图像采集和获取(包括成像方法、摄像机校正等)	58	~ 12
		A 2: 图像重建(从投影重建图像)	49	~ 10
		A 3: 图像变换,滤波,增强,恢复或复原等	86	~ 17
		A 4: 图像压缩编码(包括算法研究,国际标准实现等)	238	~ 47
B: 图像分析	433	B 1: 边缘检测,图像分割	187	~ 37
		B 2: 目标表达,描述,测量(包括二值图处理,数学形态学等)	84	~ 17
		B 3: 目标形状,纹理,空间,运动等的分析	50	10
		B 4: (2-D)目标识别,分类和提取	112	~ 22
C: 图像理解	150	C1: (序列、立体)图像(特征点)配准和匹配	71	~ 14
		C2: 3-D 建模,客观场景恢复	64	~ 13
		C3: 图像解释,推理	15	~ 3
D: 技术应用	305	D 1: 硬件,(硬件)系统	65	13
		D 2: 视频,通信	40	8
		D 3: 文档(包括文字,数字,符号等)	57	~ 11
		D 4: 生物医学	56	~ 11
		D 5: 遥感	51	~ 10
		D 6: 其它	36	~ 7
E: 综述评论	14	E1: 综述(介绍图像处理、分析、理解或它们的综合进展)	14	~ 3

### 1.3 中国图像工程的 4 个研究新热点

科学研究的成果在相当程度上体现在学术论文的发表上。通常,一门学科的重要期刊上所刊载的学术论文水平较

高,能够较好地反映该学科的最新研究成果、进展以及前沿动态<sup>[11]</sup>。所以,通过对有关图像工程重要期刊上刊载论文的统计分析,可以帮助人们了解我国图像工程研究应用的总体

情况和热点所在

进入新世纪, 图像工程也有许多新的进展。随着知识产权保护、智能化的人机交互、多传感器信息融合和高层信息处理应用等得到广泛的重视和研究, 一些新的图像工程研究热点逐渐形成。为此, 在对中国图像工程 2000 年文献的综述中, 增加了三个新的小类, 使总的小类数达到 21 个<sup>[7]</sup>。另外, 在原有的小类中, 有关图像匹配小类的文献数量在这两年有了快速的增长, 已从前 5 年在 18 个小类中平均排第 6 位升到了 2001 年在 21 个小类中排第三位<sup>[8]</sup>。这是多传感器图像的融合随着成像技术的发展和计算机运算速度的提高而得到了广泛重视的结果。由于其中的主要技术仍是图像的配准和匹配, 所以根据分类原则, 相关文献均分入 C1 小类。对这 4 个小类最近两年的文献统计情况如表 2 所示, 更全面的统计可见文献[8]。

表 2 中国图像工程 4 个新热点的文献情况

大类及名称	小类及名称	2000 年 文献 数量	2001 年 文献 数量
A: 图像处理	A 5: 图像数字水印和图像信息隐藏	20	31
B: 图像分析	B 5: 人脸和器官的检测与定位	14	15
C: 图像理解	C1: 图像配准、匹配、融合	32	36
C: 图像理解	C4: 基于内容的图像和视频检索	16	21

对照表 1 可见, 从 2000 年起新增的三个小类(A 5, B 5, C4)的文献数量均高于前 5 年小类的平均数; 由表 2 还可看出, 这些数量仍在迅速增加。这三个小类分别属于图像处理、图像分析和图像理解三个大类, 这表明在图像工程的三个层次都有了新的进展。另外, 文献数量得到较大增长的小类 C1 也属于图像理解大类, 这也印证了图像工程研究正向高层扩展的趋势。

下面对这 4 个图像工程的近期研究热点给以概括介绍

## 2 图像数字水印和图像信息隐藏

### 2.1 数字水印特点

数字水印是一种数字标记, 一般包含版权所有者的标记或代码, 以及能证实用户合法拥有数据的用户代码等基本信息, 将它秘密地内嵌到数字产品中可以帮助识别确定产品的内容、著作权、使用权、完整性等<sup>[12]</sup>。由于将水印嵌入到数字产品中, 就可将数字产品与其所有者或使用者建立对应关系, 所以也有人称数字水印技术为数字指纹技术。

近年来, 随着数字产品的广泛应用, 电子商务的全面普及, 国际标准化组织 (ISO) 和国际电信联盟 (ITU) 联合的运动图像专家组 (MPEG) 从 2000 年正式启动了国际标准 MPEG-21 的制订工作<sup>[13]</sup>。MPEG-21 称为多媒体框架, 其中知识产权保护是重要的内容。这个标准的制订推动和促进了对数字水印技术这种新型的著作权保护技术和工具 (IPM P tools) 的研究和应用<sup>[14]</sup>。

数字水印技术与一般为保密而使用的密码技术不同 (密码技术也是一种著作权保护技术), 密码技术在数据被接收及解密后就无法保护数据了, 而数字水印不仅在数据传输过程中能对数据进行保护以防止遗失或泄密, 而且在数据的整

个使用过程中都对数据使用的合法性进行保护。

数字水印的概念最初是在 1993 年提出来的, 并首先用于图像数字水印。图像数字水印是一种特殊的水印, 这与图像的可视性有关。对一般的数字产品, 可以在产品上加入可见的标志, 但对图像来说, 这样可能影响图像的视觉质量和完整性。所以图像数字水印常指隐形水印, 即水印对版权所有者是确定的, 而对一般的使用者是隐蔽的。

对图像嵌入的数字水印根据使用目的和场所的不同一般应具有以下一个或多个特性:

(1) 不可见性 也称不易察觉性, 还有人从另一个角度将其称为显著性。从人的感知角度来说, 图像数字水印的嵌入应以不使原始图像有可察觉的失真为前提。

(2) 稳健性 也称抗攻击性或鲁棒性。图像数字水印应有抵御外界处理的能力, 在多种图像处理 (如采样、量化、AD 转换、扫描、低通滤波、几何校正、压缩编码、打印等) 使得图像产生失真的情况下, 应仍保证其自身完整性和对其检测的准确性。需要注意, 在需要验证原始媒体是否被变动或破坏时可使用易损 (fragile) 水印, 易损水印对外界处理有敏感的反应。

(3) 安全性 主要指水印不易被复制和伪造的能力, 以及不易被非法检测的能力。

(4) 低复杂性 低复杂性指对水印嵌入和提取的计算复杂度低, 便于推广应用。图像由于其数据量大, 所以低的计算复杂度对水印的应用很重要。

### 2.2 图像数字水印技术原理

数字水印技术的基本原理是在拟保护的数字产品中秘密地嵌入著作权信息, 以标识真正的所有者的所有权。从信号处理的角度看, 嵌入水印的过程可看作在强背景下叠加一个弱信号的过程; 而检测水印的过程则是一个在有噪信道中检测弱信号的过程。从数字通信的角度看, 嵌入水印的过程可看作在一个在宽带信道上用扩频通信技术传输一个窄带信号的过程。

图 3 给出嵌入和检测图像水印的过程示意图。在水印嵌入端 (产品制作端), 将水印通过内嵌操作加入到原始图像中去, 就可得到嵌水印图像。嵌水印图像可通过网络或其它途径到达使用者。在水印检测端 (产品使用端), 通过对待检测图像的相关检验 (常需要结合原始水印/图像), 可判断图像中是否嵌入了水印以及所得到的判断的置信度。

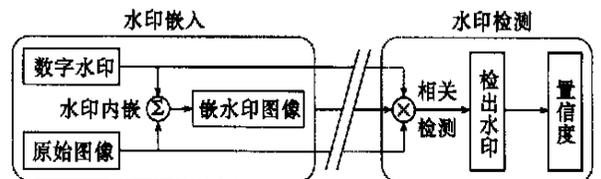


图 3 水印的嵌入和检测示意图

设原始图像为  $f(x, y)$ , 水印为  $w(x, y)$ , 嵌入水印的图像为  $g(x, y)$ , 水印嵌入过程可表示为

$$g = E(f, w) \quad (1)$$

其中  $E(\cdot)$  为嵌入函数。给出待检测图像  $h(x, y)$ , 可从中抽取待验证的可能水印  $w(x, y)$

$$w = D(f, h) \tag{2}$$

其中  $D(\cdot)$  为检测函数 若对相关函数  $C(s, t)$ , 有

$$C(W, w) > T \tag{3}$$

则认为水印存在; 否则认为水印不存在 其中  $T$  为(预先)确定的阈值

根据所嵌入水印的内容, 水印可分为无意义水印和有意义水印<sup>[15]</sup>. 前者常利用伪随机序列(Gauss 序列、二进制序列、均匀分布序列), 而后者则是可视的(文字串、印鉴、图标、图像等). 对无意义水印, 在检测时常用假设检验

无水印:

$$H_0: g - f = n \tag{4}$$

有水印:

$$H_1: g - f = w + n \tag{5}$$

用伪随机序列作为水印, 只能给出“有”和“无”两种结论, 相当于放入媒体的秘密信息为一个比特 而有意义水印可提供的信息就得多得多, 但对其嵌入和检测的要求也要高许多.

### 2.3 图像数字水印的发展与问题

图像数字水印是一个进展相当迅速的研究领域 例如, 早期图像水印的嵌入常使用空域方法, 而目前图像水印的嵌入多使用变换域方法, 它与空域方法相比, 主要优点是<sup>[16]</sup>:

- (1) 水印信号的能量可分布到所有像素上, 有利于保证不可见性;
  - (2) 可以比较方便地结合人类视觉系统的某些特性, 有利于提高稳健性;
  - (3) 变换域方法与大多数国际标准兼容, 可直接实现压缩域内的算法, 提高效率.
- 又如, 在变换域中, 图像被分解为直流(DC)分量和交流(AC)分量 传统的理论认为, 水印应尽量嵌入到AC分量中以避免改变DC分量 而现在从稳健性的角度看, 在保证水印不可见性的前提下, DC分量比AC分量更适合嵌入水印 一方面, 由于与AC系数相比, DC系数的绝对振幅大得多, 所以感觉容量要大; 另一方面, 根据信号处理理论, 嵌入水印的图像最有可能遭遇到的信号处理过程, 如压缩、低通滤波、次抽样、插值、D/A, A/D 转换等, 对DC分量的保护比AC分量要好<sup>[15]</sup>.

由表 2 的数据也可看出, 图像数字水印的研究进展相当迅速, 尽管其还是一个比较新的研究领域, 但已出现了相当数量的综述文献<sup>[7]</sup>, 例如从图像工程综述系列中就可查到文献<sup>[16-24]</sup>, 这与其它研究领域情况有所不同

当然, 图像数字水印研究所涉及的内容比较多, 还未形成比较完整的学科理论体系<sup>[17]</sup>, 值得研究的内容还很多.

首先从图像水印应具有的特性出发, 以下两方面的研究需要加强:

(1) 不可见性 不可见性是一个模糊的主观概念, 这方面的研究应结合视觉系统的特性, 然而困难是缺乏对视觉系统特性的定量描述, 那么是否可建立相应的客观标准呢?

(2) 稳健性 稳健性与隐藏信息的抗攻击性密切相关, 对水印的攻击包括降低检测水印的可能性(例如改变数据使内嵌于其中的水印标记无法辨认)和降低水印作为指示合法所有权的能力 对前者人们已建立了测试集, 以测试对水印抗攻击的能力, 但其完备性和有效性等还值得改进 对后者, 由于还涉及法律、法规等方面的问题, 需要得到更广泛的关注

另外, 图像水印的研究已从静止图像向序列图像发展, 这带来了一些新的问题 例如, 由于目前许多视频压缩方法基于对视觉上不敏感信息的去除, 因此对水印嵌入的部位有一定的限制; 但反过来, 如果将水印加入到视觉敏感部位, 对原视频信息又难以进行有效压缩 虽然这个问题对静止图像也存在, 但对序列图像这个问题严重得多 又如, 视频序列是一组密切相关的图像, 这使得水印的攻击者有可能利用图像间的互相关, 采取帧平均、帧组合的方式来消除水印 所以针对视频序列的抗攻击技术也值得研究

最后, 人们常将数字水印与信息隐藏结合讨论<sup>[25]</sup>, 但由于他们的目的不完全相同, 遇到的问题也各有特点, 因此, 所需采用的技术应充分考虑他们的应用背景和要求

## 3 人脸及器官的检测、定位和识别

### 3.1 人脸识别及相关技术的应用需求

近年来, 随着对人类行为科学、人机交互技术的研究, 以及在计算机动画设计、视频会议或远程教育系统、自动身份验证、视觉监视监控系统等应用领域的需求, 计算机人脸及器官的检测、定位和识别技术受到了广泛的重视<sup>[26]</sup>. 这方面的工作以人脸的检测和识别为典型代表, 还包括头部运动的跟踪、体态和运动信息的提取、手势手语的识别、嘴唇运动的辨识以及人脸表情的识别等等 人们试图通过这些工作, 把握人类不同情况下的行为举止, 探索友好的人机交互方式, 并满足更有个性化的服务和人际交流的需要 另外, 人脸识别在许多图像压缩编码和传输应用中也是关键的步骤<sup>[27]</sup>.

人脸识别是人类视觉的一种典型功能, 而计算机人脸识别则是利用计算机通过对人脸图像的分析、特征的提取, 进而实现(自动)辨识人脸, 并进行自动身份验证的技术<sup>[28-29]</sup>. 广义上的人脸识别是一个复杂的过程, 一个计算机人脸识别的流程如图 4 所示, 它包括了几个步骤: 对采集到的图像, 首先进行人脸的检测(在输入图像中寻找人脸), 给出人脸有无的结果; 然后进行人脸定位, 确定人脸的位置并提取出来 对人脸的定位在输入是图像序列时一般也称为人脸跟踪 通常, 检测和定位结合进行 对提取出来的人脸借助人脸描述就可进行(狭义的)人脸识别, 即通过提取特征等来确定其身份 利用不同阶段的识别结果可帮助完成不同的判断决策

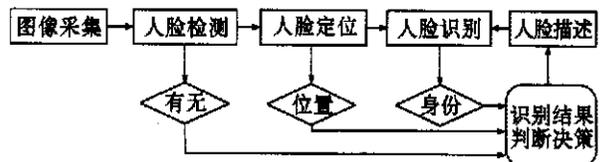


图 4 计算机人脸识别流程图

为进行人脸的检测和定位, 需要对脸部器官(如眼、鼻、嘴等)进行检测和定位; 而为进行人脸识别, 也常需要对脸部器官和特征进行识别, 所以对人脸的检测、定位和识别是与对脸部器官的检测、定位和识别密切联系的 另一方面, 不仅对人脸的检测识别技术得到了广泛的重视, 人们对人体其它部分, 人脸上各器官的检测也越来越关注, 如手势的识别<sup>[30]</sup>, 手臂的运动分析<sup>[31]</sup>, 腿部运动的跟踪<sup>[32]</sup>, 唇动的检测与定

位<sup>[33]</sup>, 面部表情的分析<sup>[34]</sup>等

### 3.2 人脸识别技术的分类和特点

人脸识别的基本思路是用知识或统计的方法对人脸建模, 然后通过比对的方法进行检测识别。如在检测定位中, 可以比较所有可能的待检测区域与人脸模型的匹配度, 从而得到可能存在人脸的区域; 又如在识别中, 可以比较提取出的人脸与人脸模型的匹配度, 从而辨别或验证所提取出的人脸的身份。建模的方法大致可分为基于统计和基于知识两类<sup>[28]</sup>。

基于统计的方法将人脸图像看作一个高维的矢量, 把对人脸检测的问题转化为对高维空间中分布信号的检测问题, 而把对人脸识别的问题转化为在高维空间中信号匹配的问题。典型的方法有主分量分析法、子空间方法、双子空间方法、局部特征法、空间匹配滤波器方法等。这类方法的主要特点是: (1) 通过对实例的训练和学习来获取需要的参数, 提高对确定性应用的可靠性; (2) 通过增加训练和学习的实例可以扩充检测模式的范围, 提高检测系统的鲁棒性; (3) 利用局部特征, 在一定程度上受遮挡等影响较小, 可用于复杂背景图像的情况; (4) 不依赖于先验知识和参数模型, 可以避免知识获取不精确或不完整而造成的错误。

基于统计的方法比较适合对正面人脸的检测, 而当人脸旋转、变换姿态或在运动中时, 问题会变得比较复杂, 效果不好。这个问题在对人体肢体等的检测识别中更为明显。另外, 统计的方法在很多情况下要对所有可能的状态进行穷举搜索, 所以计算复杂度会相当高。

基于知识的方法主要是利用人的先验知识对待识别对象建立若干规则, 从而将对人脸检测识别的问题转化为一个假设/验证的问题。这里常用的规则包括人脸轮廓规则、肤色颜色规则、目标运动规则、器官几何规则等, 还可以包括启发式的规则。目前用得比较多的是基于人脸五官的空间分布规则。例如, 可取人脸上的眼睛、鼻子、嘴巴、下巴等部件的形状和结构关系进行几何描述建立模型。基于知识的方法不需要进行大量的统计分析, 一旦规则建立后, 可以比较快速地完成检测识别任务。但基于知识的方法要建立通用规则比较困难, 规则通常仅对特定应用比较合适。另外, 规则的复杂性也不好确定, 太简单不够用, 太复杂又难于建立。

### 3.3 人脸和器官识别的难点以及应考虑的因素

人类对人脸有较强的识别能力, 但计算机要完成类似工作却有许多困难, 所以人脸识别目前主要处于实验室研究阶段, 与完全实用化尚有一定距离。计算机自动识别的难点主要有: 成像条件环境, 如光照、距离、角度等的影响, 头发、眼镜等的遮挡影响, 成像对象的姿态、人脸随年龄增长、随表情等变化的影响。另外, 人脸是立体的, 但一般识别时的成像却是平面的, 所以这一般是一个病态问题。至于对人的肢体器官的识别, 常常还受到背景环境的影响。

要对人脸和器官进行识别常先要将待识别对象从图像中分离提取出来, 这在本质上是一个图像分割问题。图像分割本身就是一个难题, 长期以来一直得到人们的关注和研究, 并已有许多成果发表<sup>[35]</sup>。但由于尚无通用的分割理论, 现已提出的分割算法大都是针对具体问题的, 并没有一种适合于所有图像的通用分割算法。所以, 要提取待识别对象还没

有完全鲁棒的方法。另一方面, 给定一幅实际的图像, 如何从众多算法中选择一个合适的算法来提取待识别对象也还没有标准的方法。通过对图像分割进行评价来自动选取最优分割算法是一个可能的方案<sup>[36]</sup>。

对人脸特征的提取和识别可看作是一个对 3-D 物体的 2-D 投影图像进行匹配的问题。由于人脸图像在获取过程中会受到光照强度、光源方向等不确定性因素的影响, 因此如何提取和表达特征也是一个难点。目前, 特征的选取多取决于匹配方法和应用要求, 如在基于 KL 变换的特征脸方法中, 需要根据训练样本构建子空间, 用其特征矢量与待识别人脸图像在子空间的投影进行匹配比较; 又如基于弹性模型的方法中, 将人脸建模为可变形的 3-D 网格, 通过确定某些特征点可把识别问题转化为可变形曲面的弹性匹配问题。

实际的人脸(也包括人的器官)是一个 3-D 的非刚性目标, 一方面其外观具有很强的多样性, 结构复杂且表面具有丰富而细致的模式变化(从颜色、纹理、轮廓、结构等特征看都不同, 随性别和种族的不同而不同); 另一方面还可能随时间和空间等发生变化(例如年龄增长的影响, 表情变化的影响, 光照条件和干扰等)。对人脸识别的研究不仅是一个视觉计算的问题, 还与解剖学和生理学密切相关, 在一些特殊的场合还要考虑人类进行识别的心理因素。因此, 人脸表情识别是一个多学科交叉的、极具挑战性的研究课题。

根据目前的研究成果来看, 对人脸的识别要结合整体识别和特征识别。从前面对相关技术的分类讨论来看, 无论是基于全局知识进行建模的方法还是基于局部特性进行统计的方法都有一些不易克服的困难, 需借助两类方法的互补来提高识别成功率。另外还有一点值得指出, 匹配两幅静止图像是相对较容易的, 但要在实际中动态地确认某个人的身份则会困难得多, 特别是有时需要“填充”那些被遮挡住或丢失的部分, 并且借助知识感知出整个脸来则更为不易。这也是人们对识别的期望和实际效果有较大差距的重要原因。

人脸和器官的识别是一个试验性很强的研究领域, 需要搜集大量的图像。在人脸检测和定位方面, 目前尚没有专门测试算法的图像库<sup>[28]</sup>。在人脸识别的实验图像库中, 包含图像最多的美国国防部 FERET 库不能在美国以外获得, 而其它库则图像数量较少<sup>[29]</sup>。在我国建立自己有权权威性的人脸实验库和规范的测试标准非常有必要。

## 4 多传感器图像信息融合

### 4.1 融合简介

多传感器图像信息的融合是将多个不同类型传感器(如雷达、声纳、CCD 摄像机、红外探测器、激光探测器等)获取的对应同一场景的不同图像数据进行空间配准, 在此基础上将各图像数据所含的优势信息进行相互补充, 并有机结合起来产生新的、信息含量更大的图像的综合分析技术。这样得到的图像能最大限度地利用各个信息源的信息, 提高分辨率、灵敏度、作用距离、测量精度、抗干扰能力等, 弥补单一信息源的不足。与由单一传感器获得的图像相比, 可以减少对场景中目标感知的不完全性、不确定性以及误差, 提高对目标

识别的准确率、对图像的综合分析精度和场景解释的能力。例如,将两个不同时相的图像融合可发现景物的变化;而将两个不同分辨率的图像融合可获得较高的分辨率和较多的光谱信息,进而提高识别精度。

多传感器图像融合的基本流程图如图 5 所示。采集的图像经过预处理后首先要相互配准,在此基础上进行融合;对融合图像进一步后处理以用于识别和解释,根据所获得的结果可以帮助作出决策。这里图像配准是图像融合的先决条件和基础,而图像识别则是图像融合的主要目的。

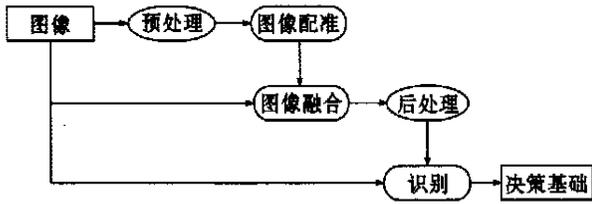


图 5 多传感器图像融合的基本流程图

国外在 70 年代就有利用多部雷达进行综合跟踪的研究<sup>[37]</sup>,和将雷达图像与卫星图像综合进行地质解释的报道<sup>[38]</sup>。80 年代以来,信息融合技术得到迅速发展<sup>[39]</sup>,并在空间开发<sup>[40]</sup>、机器人路径规划<sup>[41]</sup>、多分辨率多谱遥感成像<sup>[42-43]</sup>、医学诊断<sup>[44]</sup>等领域得到应用。在军事方面也有不少报道,如战场数据融合<sup>[45]</sup>、微波雷达和前视红外结合进行自动目标识别<sup>[46]</sup>、海军战术数据融合<sup>[47]</sup>等。近年来,随着图像采集方式的增加,计算机运算能力的提高以及许多应用领域的需求(仅军事领域就有巡航导弹路径规划和制导、无人驾驶飞机飞行控制、合成孔径雷达图像分析、战场侦察、战场地形识别、战场环境建模表示等),国内外对多传感器图像信息融合技术的研究和使用又掀起一个高潮。

### 4.2 图像配准

图像融合涉及的不同传感器获得的图像类型很多,如雷达图像、近景摄影图像、摄像机图像、航测图像、卫星图像、多光谱图像等。由于获取的方式不同、方位不同、时间不同、空间分辨率不同、空间覆盖场不同,因此要进行融合首先要将它们进行配准。图像配准是将不同图像归一化的过程,根据图像的定义和表达,需要考虑的主要有图像的采集时间、空间位置和性质。一般来说,时间归一化要求不高,性质主要由成像机理所决定,所以图像配准主要从空间上来考虑。这里因为对不同的传感器很难建立统一的数学物理模型,所以要针对具体的图像传感器建立合理的模型和采用相应的算法,如文献[48-53]。

要实现图像配准,就要对图像进行匹配,主要方法可分为两大类:区域相关和特征匹配。区域相关是比较全局的方法(区域也可以是整个图像区域),例如可直接利用模板内像素灰度的相关性,或可利用区域的矩、惯量等效椭圆<sup>[54]</sup>等。这类方法不受特征检测精度的影响,比较抗噪声,精度比较稳定,但对景物表面结构和光照条件比较敏感。

特征匹配相对来说是比较局部的方法,它利用图像中的边缘、目标的显著点等来确定配准的控制点。一旦确定了足够的控制点,对其它点进行配准几何变换的参数就可计算出来。两种方法也可结合,如可将图像特征提取与基于松弛法

的整体图像匹配结合,自动地获取密集的配准点对作为控制点,并进而组成密集三角网实现图像的精确配准。

进行图像配准的几何变换参数需根据控制点来计算,所以控制点的精确度对最终配准的精确度会有很大的影响。实际中,常需将对控制点的检测精度提高到像素内部,即达到亚像素级<sup>[55-57]</sup>,这样在图像融合时可实现较高的精度<sup>[58-59]</sup>。

### 4.3 信息融合层次和模型

图像信息融合按信息抽象程度的不同(也对应完成不同级别的功能)可分为三个从低到高的层次:像素级(原始数据)融合、特征级(或目标级)融合、决策级融合<sup>[60]</sup>。图 6 给出三个层次的特点和关系。图像融合从配准的图像出发,经过特征提取、属性判决而得到融合结果。像素级融合在低层次上进行,直接对配准后的不同传感器图像数据进行融合,根据融合的结果提取特征。特征级融合在中间层次上进行,对特征提取后获得的特征信息(如边缘、轮廓等)进行融合,根据融合的结果描述属性。决策级融合在高层次上进行,对目标属性的描述信息进行融合,根据融合的结果解释图像。上述三个层次与图像工程的三个层次有一定的对应关系。在实际中要根据需要选择和结合不同层次融合的特点,获得全局最优的效果。



图 6 图像融合的三个层次

像素级融合的准确性最高,能够提供其它层次上的融合所不具备的细节信息,但因其需要处理的数据量大,所以对于计算机的运算速度和内存容量要求较高。决策级的融合在高层次上进行,所需要处理的数据量小,且比较适合于多类不同传感器,但它由于比较抽象和模糊,精度可能较差。特征级的融合介于两者之间。

在上述每个层次上还可以使用不同的结构模型,包括不同的策略、不同的模块、不同的数据流向<sup>[61]</sup>。例如,在像素级可使用(1)分散式结构(数据的处理按各个通道分别进行)、(2)并行结构(将各个通道处理的数据统一结合起来)、(3)串行结构(各个通道串行,每次结合处理前面通道的结果和本通道的数据)和树状结构((2)和(3)两种方式的组合)。又如,在特征级可使用(1)集中式结构(集中处理多目标的信息)、(2)分布式结构(分别同步处理各个目标的信息)、(3)混合式结构((1)和(2)两种方式的并行组合)、(4)多级式结构((1)和(2)两种方式的串行组合)。

### 4.4 焦点和趋势

有如下几个研究焦点和发展趋势值得指出:

(1) 图像配准方法的研究仍在继续深入。在已提出了许多确定控制点的技术,包括利用网格相关匹配、区域重心、封闭轮廓的形心、非封闭边缘的曲率极大点等方法的基础上,亚像素级的配准也得到深入研究。

(2) 图像融合方法正在更多地结合新的理论和工具。已

提出的图像融合方法主要可分为 HSI 变换融合法、小波变换融合法、分区变换融合法、主分量分析法、高通滤波法等, 而借助理性特征的融合法<sup>[62]</sup>, 借助神经网络的融合法<sup>[63-64]</sup> 等还在不断提出

(3) 融合层次逐渐提高 与图像技术总的发展趋势相对应, 图像信息融合也逐渐向高层发展, 决策级的融合和基于融合结果的目标识别、场景解释、行为规划得到了广泛关注

(4) 应用领域更加广泛 多传感器融合的应用已扩展到许多领域, 所要满足的要求也越来越多。如由重叠图像重构场景<sup>[65]</sup>, 保持 TM 图像和中等分辨率成像光谱仪器时空和频谱特征的融合<sup>[66]</sup>, 保持光谱平衡的融合<sup>[67]</sup>, 保持信息的融合<sup>[68]</sup>, 利用融合来增强散焦图像<sup>[69]</sup> 等

## 5 基于内容的图像和视频检索

### 5.1 框架和模块

视觉信息检索是信息技术的一个新的重要研究内容。视觉信息检索是复杂的, 因为与图像或视频相联系的信息数据类别很多, 既包括与图像或视频内容不直接相关但有某种联系的数据(常称为内容无关的元数据), 如格式、作者名、日期、所有权等, 也包括与图像内容相关的视觉信息的数据。后一类数据又可分为两部分: 一部分是低层或中层特征的数据, 如颜色、纹理、形状、空间联系、运动等, 以及它们的组合(常称为与内容相关的元数据), 一般来说, 这种数据与感觉因素有关; 另一部分是高层内容语义的数据(常称为内容描述元数据), 它关心图像实体和客观世界实体的关系, 或者与视觉符号和场景相联系的时间事件、感受和意图的联系。由此可见, 要实现视觉信息的有效检索, 需要基于视觉信息的内容来进行。

基于视觉内容的图像查询检索是一种重要的多媒体信息处理技术, 它的目的是从视觉数据库中提取与一个查询内容相关的图像或图像序列。一般来说, 基于视觉内容的图像查询检索系统是介于信息用户和视觉内容数据库之间的一种信息服务系统。用户通过它可按自然的方式从库中提取满足所需要内容的图像或图像序列。这类系统的通用框架如图 7 所示, 主要由 5 个模块组成<sup>[70]</sup>:

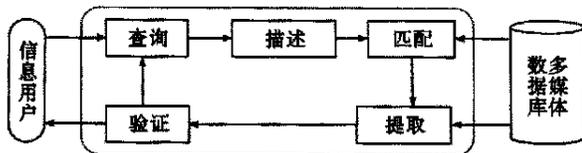


图 7 基于内容的多媒体信息检索系统框图

(1) 查询模块 对用户提供的多样的查询手段, 以支持用户根据不同应用进行各种类型的查询工作

(2) 描述模块 系统需要将用户的查询要求转化为对图像内容比较抽象的内部表达和描述, 即以一定的、计算机可以方便表达的数据结构描述图像给定的内容

(3) 匹配模块 对被查询图像建立的表达描述也要对图像库中的图建立, 这样就可以将查询描述与图像库中被查询的图进行内容匹配和比较, 以确定它们在内容上的一致性和

相似性

(4) 提取模块 根据匹配的结果, 在内容匹配的基础上将所有满足给定条件的图像自动地从图像库中提取出来, 让用户使用

(5) 验证模块 根据目前技术水平和设备条件, 在自动查询和提取的基础上, 用户还需有最后验证结果的手段。如果验证效果不满意, 新一轮的查询可通过修改查询条件而重新开始

### 5.2 基于特征的图像检索

近年来, 国际上已投入许多人力物力对基于内容的图像检索进行了研究, 国际标准 MPEG-7 制订的初衷也与此有关<sup>[71]</sup>。许多研究和应用工作集中在建立对图像内容的索引和根据特定需要对图像目标的提取上。在此基础上已研制开发了若干试验系统, 并已在一些领域试用。其中在对图像内容的描述上, 主要是利用各种图像特征来进行的。这里, 特征可以是场景或物体的颜色、表面纹理、特定目标的几何形状、几个物体在空间的相互位置关系等。

颜色是图像的一种重要视觉性质, 颜色特征的定义比较明确, 抽取也相对容易, 所以已在基于内容的图像检索中得到了广泛的研究, 已提出了许多算法<sup>[72]</sup>。

纹理也是描述图像时常用的一个概念, 但目前尚无对它正式的(或者说尚无一致的)定义。对纹理的感受是与心理效果相结合的, 用语言来描述纹理很困难。直观来说, 纹理描述可提供区域的平滑、稀疏、规则性等特性, 将纹理与颜色相结合, 可以进一步限定所提取图像的类型<sup>[73]</sup>。

形状也是描述图像内容的一个重要特征, 因而也得到广泛的重视<sup>[74]</sup>。形状常与目标联系在一起, 所以形状特征可以看作是比颜色或纹理要高层一些的特征。由于从不同视角获取的图像中目标形状可能会有很大差别, 为准确进行形状匹配, 需要解决平移、尺度、旋转变换不变性的问题。例如可采用小波变换模极大值来提取目标边缘信息, 然后用不变矩来度量相似形<sup>[75]</sup>。

场景中景物间的空间关系也是图像中的重要信息, 空间关系特征比颜色或纹理要高层一些的特征。对空间关系的表达主要有两类方法: 基于目标的方法和基于关系的方法。在基于目标的方法中, 可以使用二叉树及其各种变形; 在基于关系的方法中, 要将目标表达成符号, 如果利用符号投影将符号图像简化成 1-D 表达, 则可以使用典型的 2-D 串及其各种变形来表示空间关系<sup>[76]</sup>。

以上各种特征可以结合使用, 综合的检索功能常能取得更好的效果<sup>[77]</sup>。

### 5.3 基于内容的视频检索

数字视频是多媒体信息系统中重要的数据类型, 其特点是数据量大、信息量也大。针对这些特点, 对视频的处理有许多特殊的技术<sup>[78]</sup>, 即便是对检索结果的显示和浏览也需要相应的方法<sup>[79]</sup>。

传统的视频表示方法是将视频表示为一个无结构的比特序列——视频流, 所以需要视频进行有效合理的组织, 以便利用视频的内容进行如索引、浏览、查询、检索等操作。典型的用于组织的分层结构方案将视频分成 4 层, 如图 8 所

示<sup>[80]</sup>。这 4 层是视频节目 (program) 层、情节 (episode) 层、镜头 (shot) 层和帧图像 (frame) 层,如图 8 a 所示。根据这个方

案可对视频在不同层次进行三类操作:组织、浏览和检索,如图 8 b 所示。

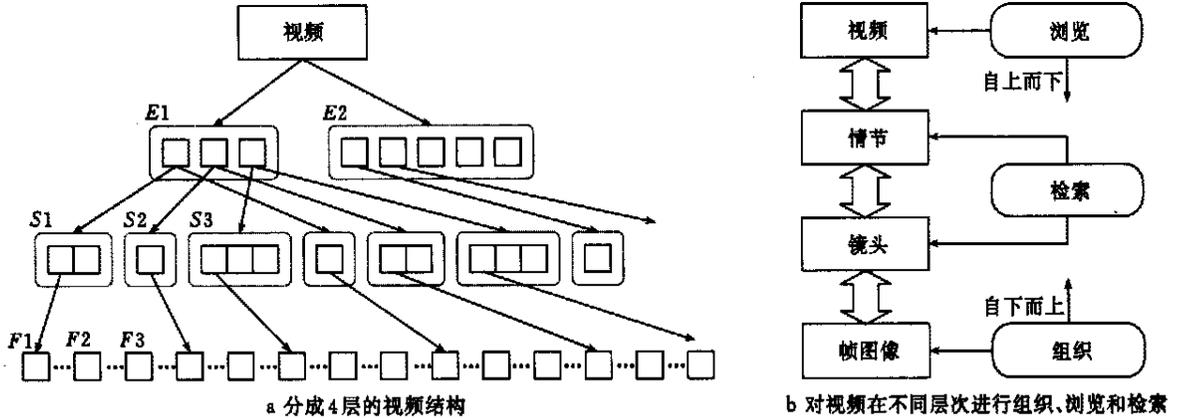


图 8 视频组织框架

为实现对视频有效合理的组织,首先需要将视频镜头转换进行检测,将时间上连续的、对应原始视频数据的帧图像组合成在相同场景拍摄的、包含空间中某个位置的一个连续动作的基本单元——镜头,这也称为时域分割。镜头切换的方式可分为突变<sup>[81]</sup>和渐变<sup>[82]</sup>,检测方法也有所不同。对每个检测出来的镜头,可以提取其关键帧作为对镜头的紧凑表达。

进一步,借助一些高层知识,可将一些镜头(不一定相连或相邻)结合成情节。情节是视频中的一种语义单元,用于描述一段故事或行动。每个视频节目都是由一系列情节构成。对每个检测出来的情节,可以提取其代表帧作为概括的表达。

视频浏览指用户在视频数据中“航行(navigate)”并发现感兴趣的视频片段/序列。一旦将一个视频节目组织成上面所讨论的结构,快速浏览就变得很直接且方便。浏览可从节目层开始由上向下进行。浏览者首先可进入情节层,发现感兴趣的情节。因为每个情节包含若干个镜头,所以进入情节后再浏览有关的镜头很容易。再进一步,镜头由一序列图像帧组成,找到感兴趣的镜头后浏览各图像帧是很直接的。

视频检索是一种直接寻找感兴趣视频序列的方法或过程。基于内容的视频检索既可在镜头层进行,也可在情节层进行;在镜头层的检索可借助镜头的关键帧进行,在情节层的检索可借助情节的代表帧进行。在这两种情况下,因为检索对象对应单帧图像,所以可使用各种基于内容对静止图像检索的具体技术来进行。

视频也可看作一系列的运动图像,所以在检索中除了可借助静止图像中的各种特征,还可以利用视频中特有的运动信息<sup>[83]</sup>。由于运动信息表达了视频图像内容沿时间轴的发展变化,所以对描述视频内容具有相当重要的作用。

### 5.4 一些新动向和进展

随着对基于内容视觉信息检索研究重视程度的增加和应用需求的迫切,有关研究内容越来越深入,涉及范围越来越广泛。例如,随着图像数据的急剧增长,图像压缩技术应用得越来越多,而从压缩后的图像中直接提取特征进行检索已有不少工作;又如,随着应用领域的扩展,人们将由具有各自

特征且互相之间具有时间和空间联系的简单目标组合成组合目标,并对具有组合目标的图像进行串行检索。此外,基于内容的图像检索技术需要随应用和使用者的不同而调整,所以可将检索过程设计成一个相关反馈的迭代过程,系统和用户通过交互逐步使检索向着接近用户期望的方向前进,并最终达到用户的要求。最后,综合利用多种媒体的特性进行检索也取得了不少有实际意义的成果,如借助音频对视频图像进行时间分割等。

尽管基于内容的视觉信息检索研究已取得了很大进展,但尚存在不少问题,离实际使用还有一定的距离。例如,目前试用的各种基于内容的图像和视频检索系统多遇到两个主要障碍:一是检索的质量常常很不一致,检索结果虽有许多需要的,但也总有许多不相关的;二是实用性比较低,如在网上检索时,很难指定一个查询图。再如当有多个特征可用时,又缺乏明确的选择依据。这些都影响了检索系统性能的提高和检索技术使用的普及。另外,虽然基于颜色、纹理、形状、空间位置等低层特征的检索方法已比较成熟,但这些低级图像特征对图像的描述与人类对图像的描述存在较大差异,所以直接利用这些特征作为检索依据常得不到满意的结果,用户也不熟悉特征值的变化对视觉效果带来的影响。

为解决这些问题,人们还在做不懈的努力。如针对仅基于低层特征检索的弊病,人们尝试利用称之为“语义特征”的高层特征。这种基于语义的图像检索技术从人类的视觉理解出发,着眼于提取图像中符合人类视觉概念的部分,与人类对图像内容的理解更为符合<sup>[84-85]</sup>。又如,为了提高系统的性能,人们引入了相关反馈技术,以充分发挥人机交互的作用。而为了适应人在检索过程中时常发生的兴趣跳转现象,还可以引入具有更大灵活性的关联反馈技术<sup>[86]</sup>。

基于内容的视觉信息检索是与对视觉信息的理解密切相关的,所以在理解基础上的检索会是一个有前途的研究领域;基于内容的视觉信息检索又是为人服务的,所以人性化的检索也会是一个值得努力的发展方向,包括如何将用户的情绪(emotion)状态或情感(affection)结合进来<sup>[87]</sup>;网络技术的发展也为基于内容视觉信息检索的研究和应用提供了契机并提出了挑战,所以,结合网络特点的相关工作值得重视。

(如由于网络上数据量更大而带来的筛选和分类工作<sup>[88]</sup>)。

## 6 结 语

本文首先对图像工程的内涵外延和中国图像工程文献综述系列进行了扼要介绍。通过对中国图像工程文献综述系列中文文献分类结果的分析,归纳出近年的4个研究新热点。结合所了解的情况和研究成果,本文对这4个研究新热点进行了全面的讨论。笔者期望该工作可以帮助人们了解我国图像工程研究应用的现状和一些特点,并能够为制定学科发展方向和研究策略提供依据材料,最终达到共同发展我国图像工程事业的目的。

## 参 考 文 献

- [1] Zhang Yujin. Image engineering in China: 1995[J]. Journal of Image and Graphics, 1996, 1(1): 78~ 83(in Chinese)  
(章毓晋. 中国图像工程: 1995 [J]. 中国图象图形学报, 1996, 1(1): 78~ 83)
- [2] Zhang Yujin. Image engineering in China: 1995(Supplement) [J]. Journal of Image and Graphics, 1996, 1(2): 170~ 174 (in Chinese)  
(章毓晋. 中国图像工程: 1995(续) [J]. 中国图象图形学报, 1996, 1(2): 170~ 174)
- [3] Zhang Yujin. Image engineering in China: 1996[J]. Journal of Image and Graphics, 1997, 2(5): 336~ 344(in Chinese)  
(章毓晋. 中国图像工程: 1996 [J]. 中国图象图形学报, 1997, 2(5): 336~ 344)
- [4] Zhang Yujin. Image engineering in China: 1997[J]. Journal of Image and Graphics, 1998, 3(5): 404~ 414(in Chinese)  
(章毓晋. 中国图像工程: 1997 [J]. 中国图象图形学报, 1998, 3(5): 404~ 414)
- [5] Zhang Yujin. Image engineering in China: 1998 [J]. Journal of Image and Graphics, 1999, 4(5): 427~ 438(in Chinese)  
(章毓晋. 中国图像工程: 1998[J]. 中国图象图形学报, 1999, 4(5): 427~ 438)
- [6] Zhang Yujin. Image engineering in China: 1999 [J]. Journal of Image and Graphics, 2000, 5A(5): 359~ 373(in Chinese)  
(章毓晋. 中国图像工程: 1999 [J]. 中国图象图形学报, 2000, 5A(5): 359~ 373)
- [7] Zhang Yujin. Image engineering in China: 2000 [J]. Journal of Image and Graphics, 2001, 6A(5): 409~ 424(in Chinese)  
(章毓晋. 中国图像工程: 2000 [J]. 中国图象图形学报, 2001, 6A(5): 409~ 424)
- [8] Zhang Yujin. Image engineering in China: 2001 [J]. Journal of Image and Graphics, 2002, 7A(5): 417~ 433(in Chinese)  
(章毓晋. 中国图像工程: 2001 [J]. 中国图象图形学报, 2002, 7A(5): 417~ 433)
- [9] Ding Xuedong. Fundamentals of Bibliometrics[M]. Beijing: Peking University Press, 1993(in Chinese)  
(丁学东. 文献计量学基础[M]. 北京: 北京大学出版社, 1993)
- [10] Zhang Yujin, Li Rui. Statistical analysis on the authors of papers cited in the survey series "image engineering in China"[J]. Journal of Image and Graphics, 2001, 6A(1): 1~ 5(in Chinese)  
(章毓晋, 李睿. 对“中国图像工程”综述系列里文献作者的统计分析[J]. 中国图象图形学报, 2001, 6A(1): 1~ 5)
- [11] Lin Beidian, Zhang Qisu. List of Core Journal in Chinese [M]. Beijing: Peking University Press, 1996(in Chinese)  
(林被甸, 张其苏. 中文核心期刊要目总览[M]. 北京: 北京大学出版社, 1996)
- [12] Hanjalic A, et al. Image and Video Databases: Restoration, Watermarking and Retrieval[M]. New York: Elsevier Science, 2000
- [13] Zhang Yujin. MPEG-21—An emerging international standard [J]. Journal of Image and Graphics, 2000, 5B(9/10): 12~ 13(in Chinese)  
(章毓晋. MPEG-21——刚开始制订的国际标准[J]. 中国图象图形学报, 2000, 5B(9/10): 12~ 13)
- [14] Li Juan, Zhang Yujin. MPEG-21 and electronic commerce [J]. Journal of Image and Graphics, 2001, 6B(7): 77~ 79 (in Chinese)  
(李娟, 章毓晋. MPEG-21 与电子商务. 中国图象图形学报, 2001, 6B(7): 77~ 79)
- [15] Zhang Y J, Chen T, Li J. Embedding watermarks into both DC and AC components of DCT [A]. In: Proceedings of SPIE, USA, 2001, 4314: 424~ 435
- [16] Huang Jiawu, Tan Tieniu. A review of invisible image watermarking [J]. Acta Automatica Sinica, 2000, 26(5): 645~ 655(in Chinese)  
(黄继武, 谭铁牛. 图像隐形水印综述[J]. 自动化学报, 2000, 26(5): 645~ 655)
- [17] Zhang Chuntian, Su Yuting, Guan Xiaokang. Digital multimedia watermarking techniques [J]. Journal of China Institute of Communications, 2000, 21(9): 46~ 52(in Chinese)  
(张春田, 苏育挺, 管晓康. 多媒体数字水印技术[J]. 通信学报, 2000, 21(9): 46~ 52)
- [18] Liu Ruizhen, Tan Tieniu. Survey of watermarking for digital image [J]. Journal of China Institute of Communications, 2000, 21(8): 39~ 48(in Chinese)  
(刘瑞祯, 谭铁牛. 数字图像水印研究综述[J]. 通信学报, 2000, 21(8): 39~ 48)
- [19] Sun Shenghe, Lu Zhen'ing. Digital watermarking techniques [J]. Acta Electronica Sinica, 2000, 28(8): 85~ 90(in Chinese)  
(孙圣和, 陆哲明. 数字水印处理技术[J]. 电子学报, 2000, 28(8): 85~ 90)
- [20] Chen Mingqi, Niu Xinxin, Yang Yixian. The research developments and applications of digital watermarking [J]. Journal of China Institute of Communications, 2001, 22(5): 71~ 79 (in Chinese)  
(陈明奇, 钮心忻, 杨义先. 数字水印的研究进展和应用[J]. 通信学报, 2001, 22(5): 71~ 79)
- [21] Hua Xiansheng, Shi Qingyun. Research on fragile watermarking problems [J]. Journal of Image and Graphics, 2001, 6A(11): 1089~ 1095(in Chinese)  
(华先胜, 石青云. 易损数字水印若干问题的研究[J]. 中国图象图形学报, 2001, 6A(11): 1089~ 1095)
- [22] Liu Ruizhen, Wang Yunhong, Tan Tieniu. Image content-based watermarking model [J]. Journal of Image and Graphics, 2001, 6A(6): 558~ 562(in Chinese)  
(刘瑞祯, 王蕴红, 谭铁牛. 基于图像内容的数字水印模型[J]. 中国图象图形学报, 2001, 6A(6): 558~ 562)
- [23] Yi Kaixiang, Shi Jiaoying, Sun Xin. Digital watermarking techniques: An introductory review [J]. Journal of Image and Graphics, 2001, 6A(2): 111~ 117(in Chinese)

- (易开祥, 石教英, 孙 鑫 数字水印技术研究进展[J]. 中国图象图形学报, 2001, 6A (2): 111~ 117)
- [24] Zhou Siqing, Yu Yinglin. Digital image watermarking and its applications[J]. Journal of Data Acquisition and Processing, 2001, 16(3): 353~ 357 (in Chinese)  
(周四清, 余英林 数字图像水印技术及其应用[J]. 数据采集与处理, 2001, 16(3): 353~ 357)
- [25] Abdulaziz N, Pang K K. Coding techniques for data hiding in images [A]. In: Proceedings of the 6th ISSPA, Malaysia, 2001. 238~ 241
- [26] Hara F, Kobayashi H. State of the art in component development for interactive communication with humans [J]. Advanced Robotics, 1997, 11(6): 585~ 604
- [27] Zhang Y J, Yao Y R, He Y. Automatic face segmentation using color cues for coding typical videophone scenes [A]. In: Proceedings of SPIE, USA, 1997, 3024: 468~ 479
- [28] Zhou Jie, Lu Chunyu, Zhang Changshui. A survey of automatic human face recognition [J]. Acta Electronica Sinica, 2000, 28(4): 102~ 106 (in Chinese)  
(周杰, 卢春雨, 张长水 人脸自动识别方法综述[J]. 电子学报, 2000, 28(4): 102~ 106)
- [29] Zhang Cuiping, Su Guangda. Human face recognition: A survey [J]. Journal of Image and Graphics, 2000, 5A (11): 885~ 894 (in Chinese)  
(张翠平, 苏光大 人脸识别技术综述[J]. 中国图象图形学报, 2000, 5A (11): 885~ 894)
- [30] Ren Haibing, Zhu Yuanxin, Xu Guangyou. Spatio-temporal appearance modeling and recognition of continuous dynamic hand gestures [J]. Chinese Journal of Computers, 2000, 23 (8): 824~ 828 (in Chinese)  
(任海兵, 祝远新, 徐光祐 连续动态手势的时空表观建模及识别[J]. 计算机学报, 2000, 23(8): 824~ 828)
- [31] Pan Chunhong, Ma Songde. 3D motion analysis of human arm based on multi-constraints fusion [J]. Journal of Image and Graphics, 2001, 6A (11): 1053~ 1057 (in Chinese)  
(潘春洪, 马颂德 基于多约束融合的人手臂三维运动分析[J]. 中国图象图形学报, 2001, 6A (11): 1053~ 1057)
- [32] Sun Yi, Chen Shunxiang, Gao Dapeng. Tracking leg joint in leg image sequence [J]. Pattern Recognition and Artificial Intelligence, 2001, 14(1): 82~ 85 (in Chinese)  
(孙 怡, 陈顺翔, 高大鹏 人体腿部运动图像的跟踪[J]. 模式识别与人工智能, 2001, 14(1): 82~ 85)
- [33] Wang Rui, Gao Wen, Ma Jiyong. An approach to robust and fast locating of lip motion [J]. Chinese Journal of Computers, 2001, 24(8): 866~ 871 (in Chinese)  
(王 瑞, 高 文, 马继涌 一种快速鲁棒的唇动检测与定位方法[J]. 计算机学报, 2001, 24(8): 866~ 871)
- [34] Pantic M, Rothkrantz L J M. Automatic analysis of facial expressions: The state of art [J]. IEEE T-PAMI, 2000, 22 (12): 1424~ 1445
- [35] Zhang Yujin. Image Segmentation [M]. Beijing: Science Press, 2001 (in Chinese)  
(章毓晋 图像分割[M]. 北京: 科学出版社, 2001)
- [36] Zhang Y J, Luo H T. Optimal selection of segmentation algorithms based on performance evaluation [J]. Optical Engineering, 2000, 39(6): 1450~ 1456
- [37] Trunk G V, Wilsor J D. Tracking filters for multiple platform radar integration [J]. Automatic Detection and Radar Data Processing, 1976, 12(1): 1~ 15
- [38] Daily M, et al. Geologic interpretation from composited radar and Landsat imagery [J]. Photogrammetric Engineering and Remote Sensing, 1979, 45(8): 1109~ 1116
- [39] Cliche F, et al. Integration of the SPOT panchromatic channel into its multispectral mode for image sharpness enhancements [J]. Photogrammetric Engineering and Remote Sensing, 1985, 51(3): 311~ 316
- [40] Rearick T C. Knowledge-based multi-sensor image fusion [J]. Lockheed Horizons, 1987(25): 22~ 30
- [41] Pau L. Knowledge representation approaches in sensor fusion [J]. Automatica, 1989, 25(2): 207~ 214
- [42] Chavez P S, et al. Comparison of three different methods to image multiresolution and multispectral data, Landsat TM and SPOT panchromatic [J]. Photogrammetric Engineering and Remote Sensing, 1991, 57(3): 265~ 303
- [43] Ehlers M. Multi-sensor image fusion techniques in remote sensing [J]. Journal of Photogrammetry and Remote Sensing, 1991, 46(1): 19~ 30
- [44] Birnbaum A, et al. Hepatic hemangias: Diagnosis with fusion of MR, CT and Tc-99m-labeled red blood cell SPECT images [J]. Radiology, 1992, 181(2): 469~ 474
- [45] Naylor R T. Battlefield data fusion [J]. IEE Conference Publication, 1987, 275: 95~ 100
- [46] Woolett J F. MMWR/FLIR/ATR sensor fusion: Proof of concept [J]. IEEE Aerospace and Electronic Systems Magazine, 1988, 3(6): 22~ 25
- [47] Byrne C D. Tactical data fusion: An application of AI to naval command and control [J]. IEE Colloquium, 1987, 49 (1): 1~ 10
- [48] Luo Shuqian, Li Xiang. Multimodality medical image registration based on maximization of mutual information [J]. Journal of Image and Graphics, 2000, 5A (7): 551~ 558 (in Chinese)  
(罗述谦, 李 响 基于最大互信息的多模医学图像配准 [J]. 中国图象图形学报, 2000, 5A (7): 551~ 558)
- [49] Yu Xiulan, Liu Shaolong, Qian Guohui. A new method for multispectral and SAR remote sensing image registration [J]. Journal of Image and Graphics, 2000, 5A (2): 100~ 105 (in Chinese)  
(于秀兰, 刘绍龙, 钱国蕙 一种多光谱和 SAR 遥感图像的配准方法 [J]. 中国图象图形学报, 2000, 5A (2): 100~ 105)
- [50] Yu Lifeng, Zu Donglin, Wang Weidong. A new method based on contour feature for multimodality medical image registration [J]. Computerized Tomography Theory and Applications, 2000, 9(1): 1~ 7 (in Chinese)  
(余立锋, 祖栋林, 王卫东 多模态医学图像的 SVD-ICP 配准方法 [J]. CT 理论与应用研究, 2000, 9(1): 1~ 7)
- [51] Cheng Fang, Zhu Minhui, Wu Yirong. Adjusted polynomial method for image registration [J]. Journal of Electronics and Information Technology, 2001, 23(11): 1071~ 1076 (in Chinese)  
(程 芳, 朱敏慧, 吴一戎 改进的多项式图像配准方法 [J]. 电子与信息学报, 2001, 23(11): 1071~ 1076)
- [52] Ge Yun, Shu Huazhong, Luo Limin. A new way of linear registration using the Legendre orthogonal moment and application in 2-value image data [J]. Acta Electronica Sinica, 2001, 29(1): 54~ 56 (in Chinese)  
(葛 云, 舒华忠, 罗立民 基于 Legendre 正交矩的配准方法及其在二值图像配准中的应用 [J]. 电子学报, 2001, 29(1): 54~ 56)

- [53] Li Zhen, Fan Xiangtao, Shi Jiancheng. Registration between remote sensing image base on multiwindow cross-correlation [J]. *Journal of Image and Graphics*, 2001, 6A (2): 129~ 132 (in Chinese)  
(李震, 范湘涛, 施建成. 基于分窗口相关的遥感图像配准方法[J]. *中国图象图形学报*, 2001, 6A (2): 129~ 132)
- [54] Zhang Yujin. Ellipse matching and its application to 3-D registration of serial cell images [J]. *Journal of Image and Graphics*, 1997, 2(8/9): 574~ 577 (in Chinese)  
(章毓晋. 椭圆匹配法及其在序列细胞图像 3-D 配准中的应用[J]. *中国图象图形学报*, 1997, 2(8/9): 574~ 577)
- [55] Tian Q, Huhns M N. A Algorithms for subpixel registration [J]. *Computer Vision, Graphics and Image Processing*, 1986, 35(2): 220~ 233
- [56] Jensen K, Anastassiou D. Subpixel edge localization and the interpolation of still images[J]. *IEEE-IP*, 1995, 4(3): 285~ 295
- [57] Zhang Yujin, Fu Zhuo. Subpixel edge detection using information along tangent direction[J]. *Pattern Recognition and Artificial Intelligence*, 1997, 10(1): 83~ 88 (in Chinese)  
(章毓晋, 傅卓. 利用切线方向信息检测亚像素边缘[J]. *模式识别与人工智能*, 1997, 10(1): 83~ 88)
- [58] Zhang Bin, Zhu Zhengzhong, Ge Chenghui. A subpixel multisensor image registration method based on edge and contour information [J]. *Journal of Image and Graphics*, 2000, 5A (10): 830~ 835 (in Chinese)  
(张斌, 朱正中, 葛成辉. 一种基于边缘轮廓信息的多源遥感图像子像素匹配定位方法[J]. *中国图象图形学报*, 2000, 5A (10): 830~ 835)
- [59] Xiong Xinghua, Qian Zengbo, Wang Renxiang. A remote sensing image subpixel matching combined genetic algorithm with least square matching [J]. *Acta Geodaetica et Cartographica Sinica*, 2001, 30(1): 54~ 59 (in Chinese)  
(熊兴华, 钱曾波, 王任享. 遗传算法与最小二乘法相结合的遥感图像子像素匹配[J]. *测绘学报*, 2001, 30(1): 54~ 59)
- [60] Solberg S, et al. Multisource classification of remotely sensed data: Fusion of Landsat TM and SAR images [J]. *IEEE T-GRS*, 1994, 32(4): 768~ 777
- [61] He You, Peng Yingning, Lu Dajin. Survey of multisensor data fusion models [J]. *Journal of Tsinghua University (Science and Technology)*, 1996, 36(9): 14~ 20 (in Chinese)  
(何友, 彭应宁, 陆大金. 多传感器数据融合模型综述[J]. *清华大学学报*, 1996, 36(9): 14~ 20)
- [62] Barron D R, Thomas O D J. Image fusion through consideration of texture components [J]. *Electronics Letters*, 2001, 37(12): 746~ 748
- [63] Zhang Zhaoli, Sun Shenghe. Image data fusion algorithm based on the one-dimensional self-organizing neural network [J]. *Acta Electronica Sinica*, 2000, 28(9): 74~ 77 (in Chinese)  
(张兆礼, 孙圣和. 基于一维自组织神经网络的图像数据融合算法研究[J]. *电子学报*, 2000, 28(9): 74~ 77)
- [64] DelCamen V M, et al. Improvement of remotely sensed low spatial resolution images by back-propagated neural networks using data fusion techniques [J]. *International Journal of Remote Sensing*, 2001, 22(4): 629~ 642
- [65] Reichenbach S E, et al. Restoration and reconstruction from overlapping images for multi-image fusion [J]. *IEEE T-GRS*, 2001, 39(4): 769~ 780
- [66] Minghelli-Roman A, et al. Spatial resolution improvement of MERIS images by fusion with TM images [J]. *IEEE T-GRS*, 2001, 39(7): 1533~ 1536
- [67] Schetselaar E M. On preserving spectral balance in image fusion and its advantages for geological image interpretation [J]. *Photogrammetric Engineering and Remote Sensing*, 2001, 67(8): 925~ 934
- [68] Liu Xiaolong. Information restore fusion of multisensor remote sensing data [J]. *Journal of Image and Graphics*, 2001, 6A (7): 636~ 641 (in Chinese)  
(刘晓龙. 多源遥感图像信息保持型融合技术的研究 [J]. *中国图象图形学报*, 2001, 6A (7): 636~ 641)
- [69] Yoon J S, et al. Enhancement of out-of-focused images using fusion-based PSF estimation and restoration [A]. In: *Proceedings of SPIE, USA*, 2001, 4310: 819~ 829
- [70] Zhang Yujin. Content-based image querying and retrieval [J]. *Information System Engineering*, 1997(8): 35~ 36 (in Chinese)  
(章毓晋. 基于内容的图像查询检索 [J]. *信息系统工程*, 1997(8): 35~ 36)
- [71] Zhang Yujin. Questions and answers about international standard MPEG-7 [J]. *Journal of Image and Graphics*, 2000, 5B (11/12): 7~ 10 (in Chinese)  
(章毓晋. 国际标准 MPEG-7 问答 [J]. *中国图象图形学报*, 2000, 5B (11/12): 7~ 10)
- [72] Liu Zhongwei, Zhang Yujin. A comparative and analysis study of ten color feature-based image retrieval algorithms [J]. *Signal Processing*, 2000, 16(1): 79~ 84 (in Chinese)  
(刘忠伟, 章毓晋. 十种基于颜色特征的图像检索算法的比较和分析 [J]. *信号处理*, 2000, 16(1): 79~ 84)
- [73] Liu Zhongwei, Zhang Yujin. Image retrieval using both color and texture features [J]. *Journal of China Institute of Communications*, 1999, 20(5): 36~ 40 (in Chinese)  
(刘忠伟, 章毓晋. 综合利用颜色和纹理特征的图像检索 [J]. *通信学报*, 1999, 20(5): 36~ 40)
- [74] Ding Xianfeng, Wu Hong, Zhang Hongjiang. Review on shape matching [J]. *Acta Automatica Sinica*, 2001, 27(5): 678~ 694 (in Chinese)  
(丁险峰, 吴洪, 张宏江. 形状匹配综述 [J]. *自动化学报*, 2001, 27(5): 678~ 694)
- [75] Yao Yurong, Zhang Yujin. Shape-based image retrieval using wavelet and moment [J]. *Journal of Image and Graphics*, 2000, 5A (3): 206~ 210 (in Chinese)  
(姚玉荣, 章毓晋. 利用小波和矩进行基于形状的图像检索 [J]. *中国图象图形学报*, 2000, 5A (3): 206~ 210)
- [76] Chang S K, Liu S H. Picture indexing and abstraction techniques for pictorial databases [J]. *IEEE PAMI*, 1984, 6(4): 475~ 484
- [77] Zhang Yujin, Xu Yin, Liu Zhongwei, et al. A test-bed for retrieving images with extracted features [J]. *Journal of Image and Graphics*, 2001, 6A (5): 439~ 443 (in Chinese)  
(章毓晋, 徐寅, 刘忠伟, 等. 通过抽取的特征进行图像检索的算法测试平台 [J]. *中国图象图形学报*, 2001, 6A (5): 439~ 443)
- [78] Jin Hong, Zhou Yuanhua. Review of video parsing techniques for content-based video retrieval [J]. *Journal of Image and Graphics*, 2000, 5A (4): 276~ 283 (in Chinese)  
(金红, 周源华. 基于内容检索的视频处理技术 [J]. *中国图象图形学报*, 2000, 5A (4): 276~ 283)

- [79] Yu T L, Zhang Yujin. Summarizing motion contents of the video clip using moving edge overlaid frame (MEOF) [A]. In: Proceedings of SPIE, 2002, U S A, 4676: 407~ 417
- [80] Zhang Yujin, Lu Haibin. Scheme and techniques for hierarchical organization of video [J]. Engineering Science, 2000, 2 (3): 18~ 22 (in Chinese)  
(章毓晋, 陆海斌. 视频分层组织方案和技术 [J]. 中国工程科学, 2000, 2(3): 18~ 22)
- [81] Lu Haibin, Zhang Yujin. An efficient algorithm for detecting abrupt scene change in video [J]. Journal of Image and Graphics, 1999, 4(10): 805~ 810 (in Chinese)  
(陆海斌, 章毓晋. 一种高效的视频切变检测算法 [J]. 中国图象图形学报, 1999, 4(10): 805~ 810)
- [82] Lu H B, Zhang Y J, Yao Y R. Robust gradual scene change detection [A]. In: Proceedings of ICIP, Japan, 1999, 2: 304 ~ 308
- [83] Yu T L, Zhang Y J. Motion feature extraction for content-based video sequence retrieval [A]. In: Proceedings of SPIE, U S A, 2001, 4311: 378~ 388
- [84] Gao Y Y, Zhang Y J, Merzlyakov N S. Semantic-based image description model and its implementation for image retrieval [A]. In: Proceedings of ICIG, 2000 657~ 660
- [85] Wang Huifeng, Sun Zhengxing. The methods of semantics processing in content-based image retrieval [J]. Journal of Image and Graphics, 2001, 6A (10): 945~ 952 (in Chinese)  
(王惠锋, 孙正兴. 基于内容的图像检索中的语义处理方法 [J]. 中国图象图形学报, 2001, 6A (10): 945~ 952)
- [86] Xu Y, Zhang Y J. Image retrieval framework driven by association feedback with feature element evaluation built in [A]. In: Proceedings of SPIE, U S A, 2001, 4315: 118~ 129
- [87] Wang Shangfei, Chen Enhong, Li J inlong. Kanse-based image evaluation and retrieval [J]. Pattern Recognition and Artificial Intelligence, 2001, 14(3): 297~ 301 (in Chinese)  
(王上飞, 陈恩红, 李金龙. 基于感性的图像评估与检索 [J]. 模式识别与人工智能, 2001, 14(3): 297~ 301)
- [88] Li Q, Zhang Y J, Dai S Y. Image search engine with selective filtering and feature element based classification [A]. In: Proceedings of SPIE, U S A, 2002, 4672: 190~ 197