



“十一五”重点规划教材
高等学校计算机及其应用系列

数字图像模式 识别方法分析

主 编 / 唐立群 郭庆昌 李永华



 哈尔滨工程大学出版社
Harbin Engineering University Press



“十一五”重点规划教材

内容简介

高等学校计算机及其应用系列

本书是“十一五”重点规划教材，是“高等学校计算机及其应用系列”中的一本。本书主要介绍数字图像处理的基本概念、基本方法和基本应用。全书共分8章，第1章介绍数字图像处理的基本概念；第2章介绍数字图像的表示方法；第3章介绍数字图像的变换；第4章介绍数字图像的增强；第5章介绍数字图像的复原；第6章介绍数字图像的分割；第7章介绍数字图像的识别；第8章介绍数字图像处理的应用。本书可作为高等院校计算机专业及相关专业的教材，也可供从事计算机工作的工程技术人员参考。

数字图像模式识别方法分析

主 编 / 唐立群 郭庆昌 李永华

图书在版编目(CIP)数据

数字图像模式识别方法分析 / 唐立群, 郭庆昌, 李永华主编. — 北京: 哈尔滨工程大学出版社, 2008.6
ISBN 978-7-5491-0000-0

Ⅰ. 数… Ⅱ. 唐… 郭… 李… Ⅲ. 数字图像处理—模式识别—方法—分析 Ⅳ. TP391.41

出版发行: 哈尔滨工程大学出版社
地址: 哈尔滨市南岗区
邮政编码: 150001
电话: 0451-82510326
传真: 0451-82510600
电子邮箱: hrbj@hrbuep.cn
网址: www.hrbuep.cn



哈尔滨工程大学出版社
Harbin Engineering University Press

内 容 简 介

迄今为止,模式识别研究取得了大量的成果,但由于实际系统中涉及到很多复杂的问题,面对这些问题,现有的理论和方法就显得有些不足了。本书主要介绍目前应用比较广泛的一些方法,以利于读者对数字图像模式识别的现状和未来发展方向有一个全面的了解。

本书重点介绍了变形曲线模型、小波分析理论以及关于这2种算法的改进。同时对均值移动算法、Adaboost、人工神经网络算法进行了详细的阐述。

本书可供高等院校的模式识别、控制理论与控制工程等专业的研究生作为教材或参考用书。

图书在版编目(CIP)数据

数字图像模式识别方法分析/唐立群,郭庆昌,李永华主编. —哈尔滨:
哈尔滨工程大学出版社,2008.6

ISBN 978-7-81133-308-4

I 数. … II. ①唐…②郭…③李… III. 数字图像-模式识别
IV. TN911.73

中国版本图书馆 CIP 数据核字(2008)第 112668 号

出版发行 哈尔滨工程大学出版社
社 址 哈尔滨市南岗区东大直街 124 号
邮政编码 150001
发行电话 0451-82519328
传 真 0451-82519699
经 销 新华书店
印 刷 黑龙江省教育厅印刷厂
开 本 787 mm × 1 092 mm 1/16
印 张 15.75
字 数 383 千字
版 次 2008 年 6 月第 1 版
印 次 2008 年 6 月第 1 次印刷
定 价 28.00 元

<http://press.hrbeu.edu.cn>

E-mail: heupress@hrbeu.edu.cn

前 言

数字图像模式识别方法已经成功地应用于工业、农业、国防、生物医学、气象、天文学等许多领域,如我们熟知的指纹识别、人工智能解谱,生物医学的细胞或组织分析等。尽管机器识别的水平还远远不及人脑,但随着模式识别理论以及其他相关学科的发展,可以预言,它的功能将会越来越强,应用也会越来越广。

本书具有如下特点:

1. 本书涉及到的方法比较新,便于感兴趣的读者掌握最新的知识动态。

2. 本书的方法涉及到很多相关领域,知识比较全面。

本书由唐立群、郭庆昌、李永华等三位老师编写。其中第五章、第六章、第七章由郭庆昌老师编写,第三章和第一章部分由李永华老师编写,其余各部分由唐立群老师编写,并负责全书的统稿。此外,曲立平、刘金路、王兴梅等多位老师在百忙之中也提出了宝贵意见。

由于编者水平有限,加之时间仓促,书中难免有错漏之处,敬请广大读者批评指正,在此表示感谢!

编 者

2008 年 6 月

目 录

第1章 绪 论	1
1.1 引 言	1
1.2 国内外研究现状	4
第2章 参数活动轮廓模型的图像分割方法研究	16
2.1 参数活动轮廓模型(Snake)的基本原理	16
2.2 基于 Snake 模型的三种经典改进算法	19
2.3 基于 Snake 模型的改进算法	25
第3章 基于水平集方法的图像分割方法研究	71
3.1 水平集理论简介	71
3.2 曲线演化理论	71
第4章 基于偏微分方程和小波变换去噪方法研究	128
4.1 小波阈值的图像去噪方法	128
4.2 基于偏微分方程扩散的图像降噪方法	131
4.3 偏微分方程与小波变换相结合的图像降噪算法	139
4.4 基于直方图和小波变换的图像分割技术	145
第5章 Mean - Shift 算法分析	150
5.1 均值移动算法	150
5.2 改进均值移动算法	173
5.3 均值移动算法应用	174
第6章 AdaBoost 算法	184
6.1 AdaBoost 基本原理	184
6.2 基于 Cascade 算法的检测框架	197
第7章 神经网络	205
7.1 人工神经元模型	205
7.2 人工神经网络学习算法	206
7.3 反向传播算法(BP 法)	207
7.4 径向基神经网络(RBF)	212
7.5 自组织神经网络(SOM)	216
7.6 Hopfield 神经网络	219
7.7 细胞神经网络(CNN)	224
参 考 文 献	241

第1章 绪论

1.1 引言

1.1.1 模式识别

20 世纪 40 年代随着计算机的出现,模式识别在 60 年代迅速发展成为一门学科。模式识别研究的理论和方法在很多科学技术领域中受到了广泛的关注,推动了人工智能系统的发展,增加了计算机应用的可能性。几十年来,模式识别研究取得了大量的成果,也成功地实现了很多应用,但由于实际系统中涉及到很多复杂的问题,面对这些问题,现有的理论和方法就显得有些不足了。本书主要介绍了目前应用比较广泛的一些方法,以利于读者对数字图像模式识别的现状和未来发展有一个全面的了解。

1. 模式识别的定义

“模式”(Pattern)这个词的原意是指供模仿用的,完美无缺的标本,这是一个相当含蓄的定义,并且触及了一些相当深奥的论题。它使我们想起了 Plato 的观点:客观世界本身正是完美理想的不尽完美的复制品。在心理学中,模式识别是作为一个过程来定义的,通过这一过程,到达感觉器官的外界信号被转换成有意义的感性经验。然而,要精确地定义什么是“有意义的感性经验”,本身就是一件很麻烦的事。

在日常生活中我们经常进行模式识别的活动。比如说,我们能够分辨出桌子、椅子,很小的时候就能够分出自己的父母,能够听出是谁的声音,能够进行正常的阅读,这些都是我们习以为常的能力,在计算机出现以前,没有人对此表现出惊奇,也没有人注意到人类的模式识别能力是一个值得研究的课题。当计算机出现以后,人工智能开始发展,模式识别也随之成为一个热门课题,当科学家发现用机器实现人类的模式识别能力是如此麻烦时,人们才意识到这个问题的难度。

一个主要的问题是模式这一概念是如何形成的,它可以由演绎或者归纳过程得到。首先,我们假定模式这一概念是观察者本身所固有的,或者假定观察者是通过许多不完全例子的观察而抽象出这一概念的。当这些被观察的例子被标以一种或几种已给定的模式时,这一过程可以称为“有导师”的学习。没有这样的标记也同样行之有效的过程,称之为“无导师”的学习。

完成这一工作的精神过程显然很复杂,且还不太为我们所知晓。模式识别是一个不仅被人类也被动物所履行的过程。并且,在进化的意义上,这个过程有着确定的生存价值。“抽象化”或者“理想化”也确实允许一个生物以一种相似于由以前的经验所证实行之有效的方法去应付新的同类型情况。较低等的生物所具备的抽象概念也许只限于危险、食物和交配,而较高级生物所具有的显然更为丰富。模式识别在生物学意义上的重要性表明,人类和动物的神经系统可能已经发展了行之有效的回路。我们也许能概括这一观察,而说神经系统对处理具备生存价值的任务要比没有生存价值的任务更有效,这样,我们就不会对人类



很容易地完成一项复杂的模式识别任务,却对做多位数乘法感到相当困难这一点感到惊奇了。

30年前,数字计算机出现了,它强有力的数值计算能力,使得人们期望它也能像人脑一样,具备非数值计算能力,然而,人们早期的乐观期望并没有成为现实,比起先进的计算机来,人脑的机制更为复杂,计算机在数值计算能力方面超过人类并不代表它的通用计算能力也很强,这种非数值计算能力如何在计算机上实现,恰恰成为人工智能和机器模式识别的重要任务。

简单一点说,模式被理解成取自世界有限部分的单一样本的被测量值的综合;模式识别就是试图确定一个样本的类别属性,即把某一样本归属于多个类型中的某一个类型。

我们这里要研究的是机器的自动识别。这就需要把人们的知识和经验教给机器,为机器制定一些规则和方法,并且让机器具有综合分析和自动分类的判断能力,以便使机器能够完成自动识别的任务,当机器学会自动识别后,可以完成人们自己难以完成的许多工作。因此,模式识别技术被广泛地应用于人工智能、计算机工程、机器人学、神经生物学、医学、侦查学以及高能物理、考古学、地质勘探、宇航科学和武器技术等许多重要的领域,随着高科技的迅速发展,模式识别技术必将获得更广泛的应用,其基本理论和方法也会愈益丰富。

2. 模式识别系统的构成

有两种基本的模式识别方法,即统计模式识别方法和结构(句法)模式识别方法,与此相应的模式识别系统都是有二个过程(设计与实现)所组成。“设计”是指用一定数量的样本(训练集/学习集)进行分类器的设计。“实现”是指用所设计的分类器对待识别的样本进行分类决策。本书的例子主要是用统计模式识别方法,在用结构模式识别的方法时,我们会对其再加以介绍。基于统计模式识别方法的系统主要由以下几个部分组成:信息获取、预处理、特征提取和选择、分类决策,如图 1.1 所示。

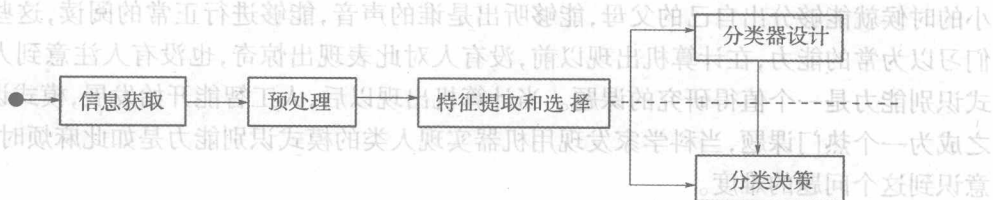


图 1.1 模式识别系统的基本构成

下面,我们对这几个部分作一下简化的说明。

(1) 信息获取

为了使计算机能够对各种现象进行分类识别,要用计算机可以运算的符号来表示所研究的对象,通常输入对象的信息有下列 3 种类型,即:

- ① 二维图像 如文字、指纹、地图、照片这类对象;
- ② 一维波形 如脑电图、心电图、机械震动波形等;
- ③ 物理参量和逻辑值 前者如在疾病诊断中病人的体温及各种化验数据等,后者如对某参量正常与否的判断或对症状有无的描述,如疼与不疼,可用逻辑值 0 和 1 表示。在引入模糊逻辑的系统中,这些值还可以包括模糊逻辑值,比如很大、大、比较大等。



通过测量、采样和量化,可以用矩阵或向量表示二维图像或一维波形,这就是信息获取的过程。

(2) 预处理

预处理的目的是去除噪声,加强有用的信息,并对输入测量仪器或其他因素所造成的退化现象进行复原。

(3) 特征提取和选择

由图像或波形所获得的数据量是相当大的。例如,一个文字图像可以有几千个数据,一个心电图波形也可能有几千个数据,一个卫星遥感图像的数据量就更大。为了有效地实现分类识别,就要对原始数据进行变换,得到最能反映分类本质的特征。这就是特征提取和选择的过程。在一些最棘手的模式识别问题中,数据以图片形式出现。由于人类视觉的物理基础尚未被很好地认识,这一课题对模式识别提出了特殊的挑战。光学错觉的存在提示我们,我们正着手于一个高级的复杂过程,并且若不考虑更高级的精神过程,视觉信息的机器处理也许是不妥当的。概括的图片模式识别将涉及下列各步:

① 图片的取样与量化 一幅景物或一张照片被转换为一个数组,以便适合计算机处理;

② 图片分割 按亮度、彩色或纹理的一致与否确定区域;

③ 景物分析 由分割获得的区域被合并或修改,使之可被定为物体;

④ 形状描述 物体被编码为反映它们形状的定量的结构;

⑤ 物体描述 可能是简单分类(例如,所见物体被分入字母A类),或者是一段语言的描述(例如,所见物体由二个被水平线连接的圆盘所组成)。

在讨论图片时总假设有一个图片文件(.bmp)存在,且可以通过函数将之读取到二维数组中。图片分割位于图像处理和模式识别的分界线,所以本书予以讨论。有一点是显而易见的,即至少在人进行模式识别时,上述各步是相互作用的。如果初始分割结果得出了“陌生”的形式,那么可能进行一个新的分割。用技术的术语来说,在分割时当我们将该物体没有先验知识时,应用自底向上的方法分割,而在有这类知识的情况下,常用自顶向下的方法。

3. 分类决策

分类决策就是在特征空间中用统计方法把被识别对象到某一类别。基本做法是在样本训练集基础上确定某个判决规则,使按这种判决规则对被识别对象进行分类所造成的错误识别率最小或引起的损失最小。

1.1.2 图像处理

自从人类发明计算机以来,人们就梦想着有一天计算机能实现像人一样的智能化。视觉是人类最高级的感觉器官,所以,毫无疑问图像在人类感知中扮演着最重要的角色。因此,数字图像处理学是实现机器智能化的一个重要方面。数字图像处理领域包括图像增强、图像复原、图像压缩和图像分割等。

一般说来,由于图像在获取、传输和存贮的过程中不可避免地受到各种噪声源的干扰,这些噪声的存在使得所观测到的图像模糊不清,该注意的细节被忽略,该识别的目标变得不可识别,严重影响了图像的应用效果。研究表明,一张模拟图像,当信噪比低于14:2分贝时,在参数估计中,参数的估计误差大于0.6%;在图像分割时产生的误检概率大于0.5%;在区域数的确定上,容易过估区域的数目。在图像的非监督分割中,区域数的准确确定对分割性能产生重要的影响,如果估计图像的区域数过少,在分割中不同的区域不能很好地分



离;如果估计图像的区域数过多,那么具有一致属性的区域可能被分割成许多相同属性的小区域,这对后面的图像处理将产生非常不利的影响,因此必须对图像进行降噪。

图像降噪指的是利用各种滤波模型,通过多点平滑等方法在已知的含有噪声的图像中去掉噪声成分。图像降噪包含两个方面,消除噪声和增强图像特征,但这两个目标在一定程度上是一对矛盾。因为去除噪声意味着除去图像的高频部分,而图像的边界也是图像的高频部分,所以在去除噪声的同时,往往使图像的边界变得模糊。如何解决好这一对矛盾是评价图像降噪模型好坏的一个重要标准。由于小波分析方法具有良好的时频局部化特性,在图像降噪时既能去除噪声又能保留边缘的信息,因此有很好的降噪效果。目前,小波分析方法在图像降噪方面取得了许多成果。

图像分割是一种基本的计算机视觉技术,是从图像处理到图像分析的关键步骤。有效、合理的图像分割能够为基于内容的图像检索、对象分析等抽象出十分有用的信息,从而使得更高层次的图像理解成为可能。目前,图像分割在医学、交通、工业自动化、文档图像处理、视频图像处理、军事以及体育等方面已经得到了广泛的应用。例如在医学应用中,将脑部图像分割成灰质、白质等脑组织和其他非脑组织区域,对后续的医学图像处理,如三维重建等有重要的作用;在交通监控的视频图像分析中,可以将车辆从目标中分割出来。总之图像分割的好坏,将直接影响到后续图像处理的质量。

图像分割方法根据使用知识的特点和层次可分为数据驱动和模型驱动两大类型。数据驱动图像分割直接对当前图像数据进行操作,虽然也使用有关先验知识,但是不依赖于知识,如 Roberts 算子、Sobel 算子等;模型驱动图像分割则直接建立在先验知识的基础上,如基于变形曲线模型的图像分割方法,以待分割目标边界是封闭的、不存在断裂为先验知识。一个好的图像分割算法应该具有以下特点:①整体性,即能得到图像有用的区域或目标的封闭边界;②精确性,即得到的区域或边界与期望的区域或目标边界很贴近;③有效性,即能将图像中有用的区域或目标分割出来;④全自动化,图像分割过程可不需要人工干预;⑤稳定性,即受噪声影响很小。

一种具体的图像分割方法很难同时满足上述各项要求,由于我们只能用图像信息中的某些部分特征去分割图像区域,所以各种图像分割方法必然带有局限性,因此只能针对各种实际情况来选择分割方法。如在医学图像分割中,由于医学图像的形状复杂、本身的模糊、不均匀性以及噪声等情况,如果采用数据驱动分割方法很难满足图像分割的要求,不能满足对分割目标要求的区域整体性,可扩展性差,不能获取(或利用)图像的总体信息。因此,必须寻找一种新的模型驱动图像分割方法进行分割。

1.2 国内外研究现状

1.2.1 图像降噪技术发展现状

人们根据实际图像的特点、噪声的统计特征和频谱分布规律提出了各式各样的降噪方法,其中最为直观的方法是根据噪声能量一般集中于高频,而图像频谱则分布于一个有限区间的特点,采用低通滤波的方式进行降噪,例如滑动平均窗滤波器、线性滤波器等。传统的图像降噪方法,是基于噪声和信号的统计特性不同,采用低通滤波的方法。在空域中,当噪声的统计特性未知时,可以采用局部平滑算子来进行降噪,这种方法的优点是可以进行并行



处理,计算量小,缺点是降噪性能受窗口大小选取的影响。若噪声的统计特性已知,在频域中,可以采用滤波器、最小二乘滤波器进行全局降噪,使用此种方法,需要知道噪声和信号的统计模型,但在实际应用中,不知道或不可能用简单的随机过程描述图像的统计模型,并且计算量也相当大。低通滤波器是一把双刃剑,在有效地消除噪声的同时,也会使图像的边缘模糊:

近年来,小波理论因具有良好的时频局部化特性,因而在图像降噪领域里得到广泛的应用,小波变换具备以下几个方面的特性。

(1)空、频局部化特性 每一小波系数能同时表征图像在空域和频域的局部内容,是一种窗口大小固定不变,但其形状可变,空域和频域的窗口都可以改变的局部化分析方法。而傅立叶变换是一种全局变换,要么完全在空域,要么完全在频域,因此无法同时表征信号的空、频局域性质。

(2)多分辨率特性 图像经一系列小波变换可以得到不同尺度下的图像特征,能较好地刻画图像的非平稳性,而短时傅立叶变换是一种单一分辨率的信号分析方法。

(3)边缘检测特性 图像经小波变换后,边缘等奇异点的小波系数幅值较大,而平稳区域的小波系数幅值较小。因此图像的边缘可以在相应的位置用较大的小波系数来表征,在降噪时对这些边缘点有所回避,从而达到降噪和保护边缘的目的。而传统的低通滤波的方法,在降噪的同时使图像的边缘模糊。

(4)能量紧支撑性 图像经小波变换后,在小波支撑范围内,边缘等非平稳区域幅值较大的小波系数较稀疏,大多数平稳区域的小波系数的幅值趋向于零,这样可以利用信号的能量来确定闭值,从而可以对图像进行自适应阈值降噪。

(5)去相关性 图像经正交小波变换后得到的小波系数近似去相关,去相关性表明小波系数在局部区域有较弱的依赖关系。而空域在一定范围内,周围点像素的相关性较强。

(6)非高斯分布特性 图像经小波变换后得到子带的小波系数呈现出重拖尾的非高斯分布,从而可选用相应的概率分布函数来逼近小波系数的联合概率分布。

(7)小波系数的持续性 图像经小波变换后,幅值较大的小波系数沿尺度间进行传播,这样有利于形成尺度间的依赖关系,为尺度间小波系数模型的建立提供可能。

小波分析的思想来源于伸缩与平移方法。它的起源可以追溯到上个世纪初。1910年 Haar 提出了小波规范正交基,这是最早提出的小波基。1982年,法国地质物理学家 Morlet 基于群论提出了小波分析这一概念;1985年至1987年间, Meyer 和 Daubechies 对连续小波理论做了进一步研究和推广,从尺度函数出发来构造小波基函数;1988年, Mallat 将计算机视觉领域的多尺度分析思想引入到小波分析当中,提出多分辨率分析的概念,用多分辨率分析来定义小波,并将 Daubechies 推导出的紧支集正交小波基离散滤波器和金字塔算法相结合,提出了 Mallat 算法。1989年,随着小波理论的进一步发展, Mallat 提出了实现小波变换的快速算法——Mallat 算法。Mallat 算法是小波分解的快速算法,为小波分析的广泛应用做出了极大的贡献,他在小波分析中的作用和地位相当于傅立叶分析中的快速傅立叶变换。小波变换是传统傅立叶变换的继承和发展。由于小波的多分辨率分析具有良好的空间域和频率域局部化特性,对高频采用逐渐精细的时域或空域步长,可以聚焦到分析对象的任意细节,因此特别适合于图像信号这一类非平稳信号源的处理,已成为一种信号/图像处理的新手段。

近年来,小波理论得到了进一步的发展,人们构造出同时具有多种优良性质的小波。



1993 年,Steffen P 和 Heller P N 等人构造了 M -带小波;1994 年,Goodman 等人提出基于 γ 重的多分辨分析,建立了多小波的基本理论框架,并给出了样条多小波的例子。同年,Geronimo,Hardin 和 Massopust 利用分形插值,成功地构造了具有短支撑集、正交和二阶消失矩的两个尺度的函数;1992 年,Vaidyanathan P P,Vetterli M,Herley C 以及 Strang G 等人发展了滤波器组理论;1996 年,Unser M,Therenz P 和 Aldroubi A 从样条函数出发构造了一组平移正交小波基等,使得小波理论不断完善。另外,1995 年,Sweldens 提出了一种基于空间域的小波构造方法——提升技术,它是构造第二代小波(the Second Generation Wavelet)的基本工具。1998 年至 1999 年,E J Candes 和 D L Donoho 建立了脊波(Ridgelet)变换,脊波在直线特征的表示和提取中非常有效,为了进一步表示多位信号中更为普遍的曲线型奇异性,又发展出局部脊波变换和曲波(Curvelet),用多个尺度的局部直线来近似表示整条曲线,从而解决了小波处理高位奇异性带来的问题。脊波和曲波的研究才刚刚开始,但它在很多领域显示出的优良性质已经为越来越多的研究者所关注。

Mallat 是最早从事小波在信号处理中应用的研究者之一,它建立了小波变换快速算法,运用于信号和图像的分解与重构,他的又一贡献是利用信号和噪声经过小波变换之后在各尺度上的不同表现,提出了一种利用小波变换模极大值原理进行信号去噪的方法,这是小波去噪的最经典方法。Mallat 通过对小波系数的模极大值处理之后,在小波变换域内去除了由噪声对应的模极大值点,仅保留了由真实信号所对应的模极大值点,然而仅仅利用这个有限的模极大值点进行信号重构,误差是很大的。因此,基于模极大值原理进行信号去噪时,存在一个有模极大值点重构小波系数的问题。Mallat 提出的交替投影(Alternative Projections,简记 AP)算法较好地解决了这个问题。然而,AP 算法计算量很大,需要通过迭代实现,有时还不稳定。

1994 年,Xu 等人提出了一种基于空域相关性的噪声去除方法,根据信号与噪声的小波变换系数在相邻尺度之间的相关性进行滤波,该方法虽不够精确,但很直接,易于实现。在该算法的实现过程中,噪声能量的估计非常关键。Pan 等人推导出噪声能量阈值的理论计算公式,并给出了一种估计信号噪声方差的有效方法,使得空域相关滤波算法具有自适应性。

Donoho 和 Johnstone 等人于 1995 年提出了信号去噪的软阈值方法和硬阈值方法,推导出计算通用阈值的公式和基于 Stein 无偏估计的 SURE 阈值公式,并从理论上证明了该阈值是最优的。同年,Coifman 和 Donoho 提出了平移不变小波去噪法。Gao 和 Bruce 把 Donoho 的软阈值和硬阈值方法进行推广,提出了 semisoft 阈值方法,通过比较说明 semisoft 阈值方法比硬阈值方法连续性好,比软阈值方法有更好的偏差等优点。Jansen 等人于 1997 年采取 GCV(Generalized Cross Validation)估计器来估计小波阈值,从而去除图像中的相关噪声。1998 年 Dowine 和 Silverman 提出了多小波的通用阈值公式,同年 Bui 和 Chen 把平移不变小波去噪推广到多小波的情形。

国内关于小波变化图像降噪研究的也比较多,1999 年,彭玉华等给出了一种基于离散正交小波变换的图像去噪方法,该方法通过二维离散小波变换将图像投影到小波变换域,通过对小波变换系数进行阈值处理实现二维图像的去噪。2005 年王正明等从理论上分析了线性模型、简单非线性模型、复杂非线性模型、多步处理模型出现的背景和优缺点,并从空域和频域上对偏微分方程模型的去噪原理进行了分析,指出了偏微分方程去噪与小波去噪结合的途径。



总之,最近几年来小波变换去噪已取得了很多好的成果。但是目前小波降噪方法大部分仍集中于高斯分布场合的研究,而对于其他噪声分布的研究不是很多,究其原因,一是由于日常出现的噪声一般都服从或近似服从高斯分布,再者是因为非高斯分布噪声往往都有比较复杂的模型,因此用现有的数学工具很难进行理论性的描述。目前人们的注意力已经转移到运用小波变换与其他方法相结合来实现对图像的降噪。

偏微分方程方法是一种有效的图像处理技术,它在图像分割、图像重建、边缘检测、图像复原等领域得到了广泛的应用。偏微分方程法从机理上来说二维方法。许多图像处理问题能够表达为最小化某个能量泛函,进一步转化为偏微分方程求解。适当规定能量泛函中对求解的约束项,可以达到边缘保持的作用。这种方法隐含了自动标定边缘位置和取向的作用,实施各向异性平滑,因而能够取得良好的效果。尤其是 Perona 和 Malik 提出的各向异性扩散方程,使得在图像降噪的同时尽量保持边缘甚至能一定程度地增强边缘,在实际应用中取得了良好的效果。

尽管小波和偏微分方程的结合面临诸多困难,但两者各自突出的优点吸引着人们不懈地努力去寻找它们的结合点。Steidl G 和 Weickert J 证明了一维情况下一次 PM 扩散等效于一次 Haar 小波收缩,并推导了收缩函数与扩散函数之间的相互关系,随后 Mrazek P 和 Weickert J 又证明了二维情况下 PM 扩散和小波收缩的等价性。近年来,Shen 和 Strang 提出的热子(Heatlets)概念在小波和热扩散方程之间建立了简单的理论联系,为该方面研究提供了新的思路。他们把任何形式的热演化看作是“原子级”热子演化的集成。而这种集成关系往往又以一种十分简单的形式给出。Shen 还进一步地研究了小波的尺度、平移不变性、消失矩等与热子演化的关系。国内关于小波变化图像降噪的研究也比较多,2005 年王正明等从理论上分析了线性模型、简单非线性模型、复杂非线性模型、多步处理模型出现的背景和优缺点,并从空域和频域上对偏微分方程模型的去噪原理进行了分析,指出了偏微分方程去噪与小波去噪结合的途径。吴亚东等分析了非线性扩散和二维 Haar 小波收缩去噪方法之间的关系,从本质上说明了非线性扩散方法优于 Haar 小波收缩去噪方法,给出了一种基于图像全变差(TV)模型的非线性扩散与二维 Haar 小波收缩相结合的混合图像去噪算法;姜东焕等给出了林石算子定义的扩散方程的小波阈值等价形式,并在此基础上对林石算子定义的扩散系数计算公式进行了修改,将其中估计各阶导数时所用的高斯线性滤波图像改成平移不变小波非线性阈值图像,这样避免了高斯滤波引起的过度光滑和边界移动。2007 年,陈莹等分析了现有的自适应建模方法在消噪性能和计算消耗上的不足,在假定小波系数为具有未知分布参数的广义高斯分布随机变量的基础上,提出了一种基于方差的上下文局部建模方法,用于估计每个系数所对应的参数。该方法能很好地反映小波系数的局部统计特性。小波和偏微分方程关系的理论研究刚刚起步,还需要进一步的研究。本文在这方面做了一些研究,对原有的算法做了改进,在对图像降噪的同时能更好地保护边缘,增强了降噪的效果。

1.2.2 变形曲线模型

近年来,基于变形曲线模型(Deformable Model)的图像分割方法已经得到广泛的关注,这种方法可以综合各种对图像数据的解释和人们的知识,因此,更接近人类的视觉机理。一般地,变形曲线模型包括一些对待分割目标的形状和特征的先验性假设,即先验模型,并以能量函数的形式来反映这种先验知识以及曲线自身的描述。变形曲线模型是定义于图像域



的曲线或曲面,并且在与曲线或曲面自身相关的内力,以及由图像数据定义的外力的作用下移动。内力的作用是保持模型的平滑性,外力的作用则将模型靠近待分割区域的边界或者其他感兴趣的特征。通过保持模型的平滑,以及综合其他诸如目标形状等的先验信息,活动轮廓线模型不仅对图像噪声和边界间隙具有很好的鲁棒性,而且可以将图像边界元素集成为相关且一致的数学表达形式,这种方式可以很方便地应用于后续更高层次的图像信息处理。

变形曲线模型以其活动轮廓线的表达方式,分为参数活动轮廓线模型(又称 Snake 模型)和曲线演化模型。不过,尽管两种方法表达形式不同,他们所遵循的轮廓曲线运动的原理是相似的。

1. 参数活动轮廓线模型(Snake)。上世纪 70 年代初,参数活动轮廓模型算法的主要原理可追溯至此。但是,只在 Kass, Witkin 和 Terzopoulos 创造性地提出活动轮廓模型之后才得到了广泛应用,这种模型通过最小化能量函数,来拟合变形模型和图像数据。最小化能量函数的基本形式就是寻找一条参数化曲线,使得基于曲线的内部能量以及基于模型的外部能量的加权和达到最小。将活动轮廓线想像为一条具有弹性的橡皮线,则内部能量就是描述活动曲线的张力和平滑性,外部能量基于图像数据定义,并在图像的边缘形成极小值。最小化内部能量和外部能量,将产生内力和外力。内力试图收缩曲线(弹力),并保持曲线不被过度弯曲(弯曲力),而外力则吸引曲线到达目标的边界。要寻找目标边缘,首先在图像内部初始化参数曲线,使其靠近待分割的边缘,然后在内外力的影响下,逐步靠近图像中的边缘。

目前对参数活动轮廓模型的研究方向主要分为 4 类:第一类对能量函数使用各种计算方法,如有限元法、神经网络和动态规划等;第二类对外力的研究,构造新的外部能量函数,提出改进的 Snake 模型,如 Cohen 提出的气球力模型、Xu 提出的梯度矢量流模型(Gradient Vector Flow, GVF)等;第三类是提出新的轮廓曲线模型,如基于 B-样条的 Snake 模型;第四类对其他问题如迭代终止标准等的研究。

对 Snake 模型的研究关键是对第二类问题外力的研究,构造新的外部能量函数 E_{ext} ,扩大 Snake 模型的捕获区以解决对初始轮廓曲线位置敏感的问题,实现图像目标的自动分割。Terzopoulos 等人提出高斯力参数活动轮廓模型,将图像外力 F_{ext} 定义为图像灰度数据与标准差(尺度)为 σ 的二维高斯函数卷积结果的负梯度值,即高斯外力。由于高斯外力是图像灰度梯度的负值,在图像灰度变化的均匀区域,高斯外力为零,不提供任何引导轮廓曲线运动的信息,轮廓曲线在内力的作用下最终收缩为一个点,高斯外力只在图像中目标边界附近有作用,远离目标边界区域时迅速下降为零。由于高斯外力只包括图像局部信息,所以只有当初始轮廓曲线在目标边界附近时,在内、外力的作用下轮廓曲线才能收敛到实际目标边界。若使用不同尺度的图像外力可达到扩展其吸引范围的目的,同时又能精确的进行图像中目标边界的提取。首先用大尺度的图像外力对原始图像作滤波处理,滤除图像中大部分噪声,以增加图像外力的作用范围,但同时强化了虚假的目标边界。当作用于轮廓曲线的内外力处于平衡时,减小高斯函数卷积图像数据的尺度 σ ,进行精确的目标边界搜索。这种方法有效地扩大了图像外力的吸引范围,但是这种算法没有给出如何选择不同尺度高斯函数滤波器的具体方法,只能凭借经验值或通过试验确定。

在高斯力参数活动轮廓模型中,初始轮廓曲线必须靠近目标边界,否则就会得到错误的图像分割结果。为扩大高斯力参数活动轮廓模型捕获区域,减少对初始轮廓曲线位置的依



赖性,1991年Cohen在高斯力参数活动轮廓模型的基础上,提出了附加另外一种外力的气球力模型(气球力Snake模型)。气球力垂直于活动轮廓曲线且为外法线方向,使活动轮廓曲线在同质区域内,从初始位置一直向外扩张。当初始轮廓曲线全部位于目标区域内部时,能正确收敛到图像中目标的边界。当初始轮廓曲线全部位于背景区域或跨越目标和背景两个区域时,气球力模型无法使轮廓曲线收敛到图像中目标的边界。若图像中目标边界太细弱,在气球力模型中的轮廓曲线会越过目标边界,无法得到正确的图像分割结果。2000年Wu等人通过结合小波系数,在气球力Snake模型基础上,提出了有向图像力(DIF)的概念,图像经过离散小波变换后,通过判断小波系数,取得过零点的方向和模值,再通过与轮廓曲线的法线方向点乘计算出有向图像力(DIF),由于该模型在判定图像力方向时只利用到局部信息,虽然能扩大吸引范围,但是该算法有时不稳定,并且小波变换明显增加了计算时间。

1993年Cohen等人提出距离力Snake模型,外部能量函数由距离映射值定义,该方法将局部棱边检测算子和参数活动轮廓模型结合起来,若棱边检测算子能正确检验出真正的图像中目标的边界,那么轮廓曲线应该是朝着与之距离最近的图像中目标的边界运动,所以,图像中每一点与最近的图像中目标的边界点之间的距离就可表示成该点的能量,该算法与气球力Snake模型相比,扩大了初始轮廓曲线的捕获区域,但计算量比气球力模型大,且轮廓曲线无法收敛到凹型目标边界。在距离力Snake模型基础上,2004年侯志强、韩崇昭基于力场分析的Snake模型,即FFA Snake模型,归纳出了感兴趣目标真边界点与假边界点的判别标准,当轮廓曲线的能量最小时,并不认为是最终图像分割结果,而是依据归纳的判别标准作出进一步的计算和判断,只有当轮廓曲线符合力场分析所确定的真基于活动轮廓模型的图像分割方法研究边界条件之后,才认为是最终图像分割结果,从而避免了距离力Snake模型容易陷入局部极小点的不足,FFA Snake模型虽然能扩大收敛范围,但其正确性依赖于棱边检测算子。上述参数活动轮廓模型无法收敛到目标的深度凹型边界问题,一些研究者提出了局部自适应法来扩大搜索区域方案,但是效果并不理想。1998年Xu等人提出了梯度矢量流GVF Snake (gradient vector flow Snake)模型这一问题。他设计了一种GVF新的外力,由扩散方程得到的新外力在整个图像区域计算梯度场,提供了一种自然地扩大Snake模型捕获区域的新机制,将Snake模型轮廓曲线拖向物体的深度凹陷区域,由于需要求解扩散方程,其计算量较大。2003年王洪元等核磁共振图像分割研究中,针对参数活动轮廓模型存在搜索区域小的问题,对参数活动轮廓模型进行了改进,引入新的面积能量,并采用二阶段搜索以及自适应动态增减节点等策略,在一定程度上克服了传统Snake模型搜索区域小,对初始轮廓曲线位置要求高,收敛速度慢等不足,具有一定的实际应用价值。

另外,参数活动轮廓模型的缺点是不容易处理拓扑结构的变化。对于结构较为复杂的图像,需要活动轮廓模型能自动改变拓扑结构,以适应分割多个子区域的需要,而参数活动轮廓模型却很难满足这种需求,尽管近年来出现了可以改变拓扑结构的参数活动轮廓模型,然而处理过程也非常复杂。

2. 曲线演化模型

1993年Caselios和Malladi等分别独立提出曲线演化模型,可以认为是Snake模型的扩展,不同于Snake模型之处在于这种模型的轮廓曲线运动(变形)过程是基于轮廓曲线的几何度量参数(法线方向矢量,曲率等)。这样,轮廓曲线运动过程就独立于轮廓曲线的参数,因此可以自动处理拓扑结构的变化。

近十年来,水平集方法(Level set Method)的出现,极大地推动了曲线演化模型的研究,



几何曲线演化理论和水平集相结合的方法,称为曲线演化模型,是目前广为关注的一种图像分割方法。曲线演化模型的特点是非常自然地处理拓扑结构的变化,而水平集方法则以一种紧凑的方式表达了轮廓曲线演化,并且提供了稳定的数值算法。

水平集方法的特点就是通过一个高维函数曲面来表达低维的演化曲线或曲面,即将演化曲线或曲面表达为高维函数曲面的零水平集的间接表达形式,并将演化曲线或曲面的演化方程转化为高维水平集函数的演化偏微分方程,从而避免了变形曲线或曲面的参数化过程。只需在固定二维或三维网格上,利用有限差分方法求解随时间变化的演化方程。并且利用轮廓检测的方法,检测演化后的水平集函数的零水平集(Zero Level Set),就可得到演化后的曲线。由于避免了曲线的参数化,因此,变形曲线(或曲面)的拓扑结构变化变得非常自然;由于曲线的演化问题转化为函数在固定网格的变化,因此,简化了问题的求解过程。由于水平集方法将曲线演化模型的演化过程转化为以水平集函数偏微分方程表达的固定模式的数值求解问题,因此,几何变形曲线的表达形式和变形过程也标准化,从而使人们对几何变形曲线的研究主要集中在分割模型的建立,即根据要解决的实际问题,构造演化速度 V 。

为了设计合适的速度函数, Malladi 利用图像的边缘强度控制几何曲线演化的速度,使变形曲线停止在图像的边缘。显然,当活动轮廓线位于图像的边缘,则图像高斯梯度的模较大,演化速度几乎为零,曲线就停止在图像的边缘位置。不过,这种速度函数对具有较好对比度的图像分割效果不错,然而,如果目标边缘不明显或者存在缝隙,则几何活动轮廓线可能从图像的边缘泄露出去,一旦出现这种情况,演化曲线就很难再返回到正确的图像边缘了。

为了设计合适的能量函数, Mumford 和 Shah 提出了一个著名的结合图像边界和区域的分割模型,即 Mumford-Shah 模型,该模型不需要对待分割图像区域的任何先验知识,完全基于图像数据的驱动来完成分割。该模型将图像的恢复、区域分割和边缘的检测综合在一起,可以说是当前图像分割中比较完美的模型。Chan 和 Vese 提出了利用水平集方法求解简化 Mumford-Shah 模型的图像分割方法。在噪声较大,图像边缘较弱的时候,实验证明这种模型分割效果很好。Caselles 等和 Yezzi 等从能量最小化的角度,提出另外一种曲线演化模型的速度函数,称为测地活动轮廓线模型(Geodesic Active Contour Model),以求一定程度解决 Malladi 所提出的分割模型的缺陷。只有在活动轮廓线到达图像边缘附近的时候才有用,主要作用是在以图像边缘为中心的附近形成能量势阱,且边缘中心的势能最低。这样,活动轮廓线一旦进入势阱周围,将被吸入势阱,并被锁在势阱底部。因此,该方法进一步提高了基于边缘方法的稳定性。Siddiqi 等在测地活动轮廓线模型上,增加一个面积约束项,以提高变形曲线跨越图像边缘的较小缝隙的能力。然而对于较大的缝隙,该模型仍然是无能为力。

1.2.3 均值移动算法

1975 年 Fukunaga 和 Hostetler 在研究核函数的密度梯度估计时提出了均值移动算法的原型“valley-seeking procedure”,尽管效率非常出色,但并未得到学术界的注意。直到 1995 年 Y Cheng 成功地将此算法扩展至计算机视觉领域,才引起了广泛的关注。开始 Fukunaga 等认为均值移动算法是一种求最大值的最速上升法。1995 年 Y Cheng 修正了 Fukunaga 的观点,认为均值移动算法是具有自适应步长的最速上升法。2003 年 Y Cheng 等提出了基于拟牛顿法的改进均值移动算法,使之达到了超线性收敛速度,他们的方法是直接计算梯度,



用 BFGS 方法估计 Hessian 矩阵, 计算迭代步长。2000 年 Comaniciu D 等证明了均值移动算法密度随均值移动点单调递增性。2005 年 M Fashing 等证明了当核函数为常数时, 均值移动算法为边界优化算法。如果要对其进行改进, 则需要收缩边界, 找到合适的边界将耗费较多的计算量, 但他们并没有给出改进的方法和途径。2005 年李乡儒基于多重假设证明了均值移动算法的收敛性。

近年来均值移动方法在计算机视觉和图像处理领域得到了广泛的应用, 例如 Comaniciu D 等讨论了基于均值移动的目标跟踪算法; 分析了均值移动算法的运动特征和其在图像分割中的应用; 应用均值移动算法进行图像平滑; 讨论了变窗宽均值移动算法; 讨论了图像聚类中鞍点的解决方法。Bradski GR 等用均值移动算法进行人脸跟踪。Collins RT 等提出了变尺度目标跟踪算法。C Yang R 等提出了基于核的目标跟踪算法。朱胜利等用卡尔曼滤波器和均值移动算法相结合进行运动目标跟踪。贾静平, 张艳宁等提出了多自由度的均值移动目标跟踪算法。李培华提出了一种改进的均值移动跟踪算法。程伟、杨杰用均值移动算法对红外目标进行跟踪。K Deguchi 等采用粒子滤波器和均值移动算法相结合进行目标跟踪。A Bandera 等通过霍夫变换和均值移动算法相结合进行图像分割等。Yuzhong Wang 等采用自适应均值移动算法和软规则对质地图像进行非监督分割。Wang J 等采用各项异性核进行图像分割和运动目标跟踪。M 等采用高斯模糊均值移动算法进行非参数聚类。Georgescu B 等在高维空间用均值移动算法聚类分析。H Tek 等基于射线传播规律采用均值移动算法进行血管检测。刘蓉等用均值移动算法去除光谱信号中的噪声。

1.2.4 Adaboost 算法

1984 年 Kearns 第一次提出了“是否可以用许多弱分类器组建成一分类能力很强的强分类器”的概念。1989 年 Schapire 第一次提出并证明了 Boosting 算法。一年后, Freund 提出了更为有效的 Boosting 算法, 但他和 Schapire 一样, 也是没能解决 Boosting 算法在实用中遇到的问题, 即它们都要求事先知道弱学习算法学习正确率的下限, 但是这点在实际应用中很难做到。1995 年, Freund 和 Schapire 提出了 Adaboost (Adaptive Boost) 算法, 它的效率与 Freund 所提算法相近, 但是可以很容易的应用到实际中, 之所以称之为“自适应”的是因为 AdaBoost 算法整合强分类器 H 的规则不是事先确定好的, 而是由构成 H 的若干弱分类器的性能决定的。1999 年 R E Schapire 和 Y Singer 对 AdaBoost 算法作了改进, 使其从处理离散二值判定规则, 即输出为 ± 1 , 推广到处理具有连续置信度输出的弱分类器, 即值域为实数域 R 。2002 年 S Z Li 等人指出由于 AdaBoost 是一种向前搜索的贪心算法, 它并不能保证单调递增性, 即每一轮迭代后得到的新的强分类器比原有的强分类器具有更好的性能; 针对这一缺点, 他们提出了一种称作 FloatBoost 的 Boost 算法, 它可以使得学习过程满足单调递增性。但 FloatBoost 的一大缺点是其训练过程需要的计算量远远大于 AdaBoost。在同一个训练集合上, 他们分别用 AdaBoost 算法和 FloatBoost 算法学习了两个人脸检测器并比较它们的性能, 试验结果表明 FloatBoost 算法所选取的弱特征个数约为 AdaBoost 算法的 $2/3$, 且性能较优。2003 年 C Liu 和 H-Y Shum 指出 AdaBoost 存在两个未解决的问题, 即如何设计弱学习算法和如何将它们以最优的方式组合起来。针对这两个问题他们提出了一种称作 Kullback-Leibler Boosting (简计为 KLBoosting) 的学习算法。首先将高维的数据投影到线性特征上, 并且沿这些特征统计各类样本分布的一维直方图, 这些直方图近似表示了数据的概率密度特性。对于两类问题, 计算两类直方图的差异就可以衡量相应特征对于分类问题的优劣。



这种直方图间差异的度量就是 Kullback - Leibler 散度 C Kullback - Leibler Divergence, 简记为 KL 散度, 而对分类问题最优的特征就是具有最大 KL 散度的特征, 称为 KL 特征。寻找 KL 特征的 Boost 算法就是 KLBoosting。实验证明 KL 特征确实要比 Haar 特征具有更强的分类信息, KLBoosting 人脸检测器只用了 450 个 KL 特征, 而正确率高于 AdaBoost 检测器。但 KL 特征比 Haar 特征复杂得多, 因此要消耗更多的计算量, 这就使它无法达到实时运算的效果。当输入为 320×240 大小的图片时, 在一台 Pentium IV 1.8GHz PC 上 KLBoosting 检测器的速度为 400 ms 每帧, 远远慢于 P Viola 和 M Jones 的 AdaBoost 检测器。2003 年 Y Ma 等同样对 AdaBoost 进行改进, 提出了一种称为 CS - AdaBoost (Cost Sensitive AdaBoost) 的算法。在原始的 AdaBoost 算法中, 各类样本的初始权重是相同的, 即对所有类别的所有样本一视同仁。但在物体检测问题中, 各类的重要性显然是不同的, 通常我们所希望的是将所有物体模式都检测出来。这种情况下, 正例样本的重要性要大于反例样本。CS - AdaBoost 的基本想法就是, 为每个样本附加一个初始权重, 这个权重将在以后的 Boost 过程中始终伴随着该样本, 权重大的样本将得到更多的重视。当所有样本权重都取 1 时, CS - AdaBoost 就退化为原始的 AdaBoost 算法。同样是在 2003 年 R Lienhart 等人同样指出对于复杂物体的检测问题, 采用分而治之或是由粗到精 (Coarse - to - fine) 的方法会更有效。由此, 他们提出了一种树型 Boost 检测器结构。树型检测器的节点与瀑布型检测器的层一样, 都是由 AdaBoost 算法挑选弱分类器而得到的强分类器。不同之处在于, 在树型检测器的训练过程中, 每学习得到一个节点后, 都要判断是否需要将当前问题划分为若干较小的子问题, 如果需要进行划分的话, 则在每一个子问题上递归的进行新一轮学习。他们用到的进行划分的判据是当前节点的复杂度, 即当前节点所包含的特征个数。当每个节点要达到的精度 (检出率、误报率) 一定时, 节点所用的特征个数越多, 说明所关心的分类问题越复杂, 也就暗示着分而治之也许会更有效。树型检测器的判断过程就是对树的深度优先遍历, 若输入模式能够找到一条到达某个叶节点的路径, 则被判定为物体, 否则为非物体。直观上看, 树型检测器与瀑布型检测器相比, 无论在训练还是检测过程中, 都应该需要更多的计算量。但 R Lienhart 等人在唇部跟踪问题上的实验表明, 树型检测器在速度和正确率上都要优于瀑布型检测器, 这也许是因为树型检测器的每个节点都具有相对较低的复杂度的缘故。

1.2.5 人工神经网络算法

人工神经网络的研究相对其他学科来说并不长, 它始于 20 世纪 40 年代初, 在短短半个多世纪的时间里, 人工神经网络的研究经历了兴起、高潮与萧条、高潮及稳定发展的曲折道路, 大致分为五个阶段。

1. 奠基阶段

早在 20 世纪 40 年代初, 神经解剖学、神经生理学、心理学以及人脑神经元的电生理的研究等都取得了丰硕的成果。1943 年, 神经生物学家、心理学家 W S McCulloch 与青年数理逻辑学家 W Pitts 合作, 从人脑信息处理观点出发, 采用数理模型的方法研究了脑细胞的动作和结构及其生物神经元的一些基本生理特性, 这是第一个用数理语言描述脑的信息处理过程的第一个神经元模型, 即神经元的阈值元件模型, 简称 MP 模型, 他们认识到了模拟大脑可用于逻辑运行的网络有一些结点, 结点与结点之间相互联系, 构成一个简单的神经网络模型。其主要贡献在于结点的并行计算能力很强, 为计算神经行为的某些方面提供了可能性, 从而开创了神经网络的研究。这一革命性的思想, 产生了很大影响。