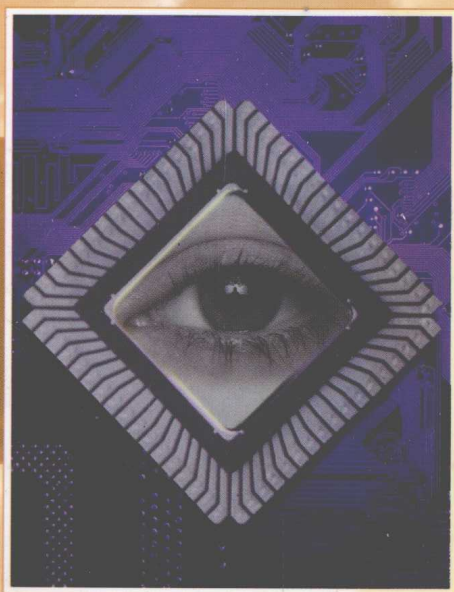


ZhiNengShiJue

Tu Xiang Chu Li

智能视觉图像处理

——多通道图像的无监督学习方法
及其他方法



[美] 斯华龄
张立明
编 著

上海科技教育出版社

智能视觉图像处理

——多通道图像的无监督学习方法及其他方法

[美]斯华龄 (乔治·华盛顿大学)

张立明 (复旦大学)

编著

上海科技教育出版社

图书在版编目(CIP)数据

智能视觉图像处理:多通道图像的无监督学习方法及其他方法/(美)斯华龄(H. Szu),张立明编著.

上海:上海科技教育出版社,2002.10

ISBN 7-5428-2994-7

I. 智... II. ①斯...②张... III. ①计算机应用—图像处理②计算机视觉 IV. ①TP391.41②TP302.7

中国版本图书馆 CIP 数据核字(2002)第 073054 号

智能视觉图像处理

——多通道图像的无监督学习方法及其他方法

编 著 [美]斯华龄

张立明

责任编辑 王福康

装帧设计 江 南

出版发行 上海科技教育出版社

(上海冠生园路 393 号 邮政编码 200235)

网 址 www.sste.com

经 销 各地新华书店

印 刷 商务印书馆上海印刷股份有限公司

开 本 850×1168 1/32

字 数 305 000

印 张 11.5

版 次 2002 年 10 月第 1 版

印 次 2002 年 10 月第 1 次印刷

印 数 1-600

书 号 ISBN 7-5428-2994-7/TP·30

定 价 40.00 元

编写本书的目的

图像处理的方法离不开人们对生物视觉的了解,如在图像压缩时,离散余弦变换(DCT)系数的量化,就是考虑到人对低频成分的敏感度比高频的强,从而对高频用大的量化步长,低频用小的量化步长,使人眼看到经过压缩的图像,仍有较高的质量.在视频采样中,用25~30帧/秒,是因为人眼不能分辨超过30帧/秒以上的图像变化.但是,多数的图像处理方法还是建立在数字信息处理和概率统计的基础上,与人的视觉处理还有很大的差别.

人工神经网络和人工智能企图模拟人脑的功能.人工智能是用模式模拟人处理事件的能力,而人工神经网络是从结构上模拟人脑的功能.它们各自都发展成一门学科,但人工智能解决的是大的决策问题,需要很多先验知识,使用符号逻辑方法,在模式识别上的应用成效较小.人工神经网络采用的是学习机理,在有监督学习中可通过“教师”来调整权,达到识别的目的,但还有很多方面没有跳出工程技术方法的范畴,在图像处理上的应用也不多.霍普菲尔德(Hopfield)的反馈式神经网络是建立在李雅普诺夫(Lyapunov)函数最小的基础上,在优化和联想记忆上有其特点,因它的连接权是通过学习获得或用目标函数进行设计的,在本书中把它归为有监督学习.细胞神经网络(CNN)主要用来进行图像处理,它的权也是预先设计好的,也归为有监督学习.关于这些方面的内容已有很多教科书作过介绍.本书讨论的内容主要是建立在无监督学习基础上,针对近年来发展很快的、以最大熵为目标的独立元分析方法,同时突出把独立元分析方法与图像和信息分析相结合的一些思路,把视频图像处理与人工神经网络相结合,形成新的、智能的、类脑的图像处理方法,给读者打开一扇新的窗口.

新的智能图像处理是一种新的模型,即有监督学习和无监督

学习并行工作,无监督学习使输出信息的熵达到最大,将获得原始信息的精锐留在神经网络的权中,而其他无用的“垃圾”被清除掉。这些权又与生物视觉中的感受野的权有一定的相似性,是伪正交的,也是稀疏的。用这些权组成的基,对输入图像进行映照,可以得到比较好的特征。一种并行工作的能量函数——亥姆霍兹(Helmholtz)自由能也被提了出来,它以温度为控制量,当自由能达到最小,使两种学习都能得到收敛。

这些新的思路为研究图像处理和神经网络的人员提供新的研究方向。为使读者更容易理解这些新的思路,本书还简要地介绍传统图像处理方法和有监督学习的神经网络方法。对独立元分析的最新几种方法在书中也作深入浅出的介绍,包括理论、算法、方法的由来等,让一些初学独立元分析的学生也能理解。书中还列举了大量独立元分析方法在图像分析、处理、识别上的应用实例。在这些实例中,大部分是应用很成功的,但也有一些应用还需要继续研究或遇到一定的困难,在书中都如实地告诉读者,以供在图像、视频及信息处理领域工作的工程技术人员进一步开展这方面研究和应用作参考。

本书的作者长期从事人工神经网络和图像处理的研究工作。斯华龄(Harold Szu)教授是美国乔治·华盛顿大学的教授、人工神经网络的先驱者。他对神经网络、图像处理以及独立元分析方法在理论上和应用上做过长期的研究,提出很多新的见解和方法。尤其在独立元分析方法的应用上,他提出一个新的目标函数——多通道、单像素无监督学习独立元分析方法,并用于医学和遥感图像处理,非常成功。他的出色研究工作得到学术界公认。张立明教授是上海复旦大学电子工程系图像与智能实验室的负责人,她和她的研究小组在图像处理和人工神经网络方面做过长期的研究,并在独立元分析方法用于图像识别方面做过一些应用性研究。本书是在斯华龄教授的积极建议下,由复旦大学图像与智能实验室的研究人员(康学雷、武研、杨代明、丁佩律、刘献容、潘逢治、梅剑锋、戴天荣、王蓓、王海川、张奇峰、梁建助等),阅读了斯华龄教授提供

的大量研究资料和有关文献,主要包括斯华龄教授和国际上几个著名的研究小组的工作,经研究人员的消化、补充,并结合他们自身的工作经验,并与斯华龄教授进行多次讨论,才编写而成的.本书的主要内容包含了图像处理、人工神经网络和独立元分析的理论及方法,将目前国际上有关智能视觉图像处理的一些前沿工作介绍给读者.

此外,周瑞华老师承担本书的编务工作,在此一并致谢.

目 录

编写本书的目的	i
第 1 章 引言	1
第 1 节 传统图像处理方法	2
1.1 图像处理	3
1.2 图像变换	16
1.3 图像压缩编码及标准	32
1.4 图像模式识别	47
1.5 图像加密(水印)	62
第 2 节 智能视觉图像处理与传统图像处理的异同	67
2.1 概述	67
2.2 人的视觉系统	69
2.3 人工神经网络模型	71
2.4 智能图像处理方法和传统图像处理方法的不同	77
第 2 章 智能视觉图像处理基础——独立元分析方法	96
第 1 节 有监督学习算法	96
1.1 有监督学习与无监督学习的联系	96
1.2 用能量函数作为目标函数的人工神经网络方法	98
第 2 节 独立元分析原理	100
2.1 独立元分析的提出	101
2.2 解独立元分析问题的发展情况	102
2.3 解的可能性分析	104
第 3 节 随机变量独立性的概念和目标函数	113
3.1 随机变量之间的独立概念	113

3.2	衡量输出独立性的目标函数	114
3.3	概率密度函数和熵的近似	123
第4节	主元分析算法	125
4.1	理论与原理	126
4.2	求解主元分析的数学方法	128
4.3	用人工神经网络训练来完成主元分析	129
第5节	基于统计平均的独立元分析算法	135
5.1	概述	135
5.2	早期的独立元分析算法	136
5.3	用目标函数的算法	140
第6节	单像素独立元分析算法	
	——拉格朗日约束神经网络	160
6.1	问题提出	160
6.2	拉格朗日约束神经网络原理	161
6.3	拉格朗日约束神经网络的迭代算法	171
6.4	讨论	172

第3章 基于统计平均的独立元分析在智能视觉图像处理

	中的应用	174
第1节	概述	174
第2节	独立元分析技术在图像获得与传输中的应用	178
2.1	概述	178
2.2	智能摄像机——用于分离多个图像的独立成分	179
2.3	通信中的应用——两条信道传送4幅图像	185
第3节	独立元分析技术在纹理分割图像上的应用	187
3.1	概述	187
3.2	独立元分析的信息集中	188
3.3	纹理数据在独立元基中的表现	190
3.4	结论	195
第4节	独立元分析技术在人脸图像识别中的应用	196

4.1	概述	196
4.2	“特征脸”方法	198
4.3	独立元分析方法用作人脸识别	200
4.4	人脸识别中的主元分析、独立元分析及局部独立元分析	203
4.5	实验结果	205
4.6	讨论	209
4.7	以后的实验方向	211
第5节	独立元分析技术在数字水印上的应用	212
5.1	概述	212
5.2	用独立元分析人工神经网络的多媒体数据盲分离	213
5.3	种入多媒体水印疫苗的独立元	215
第6节	独立元分析技术应用在语音分析上	221
第7节	独立元分析技术在某些应用中遇到的问题	225
7.1	心电信号处理	225
7.2	心脏瓣膜生物修复术后心脏异常的心音分析	233

第4章	单像素独立元分析方法在早期乳房癌检测中的应用	244
第1节	概述	244
第2节	背景和原理	244
第3节	拉格朗日约束神经网络方法和正交对消向量的方法	247
3.1	波尔兹曼熵	248
3.2	早期乳房肿瘤检测的目标函数——亥姆霍兹自由能	248
3.3	用正交对消向量的算法	250
3.4	小结	251
第4节	用亥姆霍兹自由能对早期乳房癌的检测结果	252

第 5 章 单像素独立元算法与其他方法相结合的遥感图像	
处理	255
第 1 节 概论.....	255
第 2 节 单像素的独立元分析对多光谱遥感图像的处理 ...	259
2.1 拉格朗日约束神经网络的多光谱遥感图像处理	
原理	259
2.2 拉格朗日约束神经网络算法步骤	262
2.3 实验和结果比较	263
第 3 节 集成子波纹理——边缘特征的神经网络的遥感	
图像分类器.....	266
3.1 人工神经网络分类——每个像素单个标记	267
3.2 宏观结构的子波纹理索引	267
3.3 实验结果	271
3.4 结论	273
第 4 节 超光谱图像处理.....	274
4.1 概述	275
4.2 超光谱图像处理的新方法:正交子空间投影方法	
和独立元分析方法	277
4.3 空载可视红外成像摄谱仪图像处理结果	282
第 5 节 合成孔径雷达干涉仪的应用.....	285
5.1 数字雷达	285
5.2 合成孔径雷达系统	288
5.3 用合成孔径雷达干涉仪确定三维空间	293
第 6 章 人的视觉边缘奇异图及应用	298
第 1 节 用独立元分析的初级视觉系统的图像分析.....	298
第 2 节 人的视觉系统的奇异图分析.....	304
2.1 奇异图概念	304
2.2 抽取奇异图的一般方法	306

2.3	用奇异图解决视频压缩中的光流问题	313
第3节	借助哈密顿草帽子波得到加噪信号的奇异图.....	314
第4节	奇异图和子波变换进行超高压压缩率的图像压缩	321
4.1	新的图像压缩思路	321
4.2	WaveNet 系统	322
4.3	WaveNet 算法流程	324
4.4	视频压缩模块	328
4.5	小结	329
第5节	在视频图像中对移动物体识别的约翰逊法则.....	330
5.1	约翰逊法则	331
5.2	人的视觉感知和视频图像中的约翰逊法则	331
5.3	对于“类约翰逊”法则的验证实验	334
附录:	在任意角度下的 W_{szu} 和 W_{wang} 与最大的特征 值的关系	337
后记		340
参考文献		341

第1章 引 言

图像和视频信号处理、图像信息压缩和编码、图像目标和模式识别是人的视觉处理中最为活跃的几个领域,在短短的二十年中,以其特有的魅力在所有与成像有关的领域中取得了不同程度的成功,并且由于新技术的不断出现和成熟,在可预见的将来还将得到飞速的发展.作为交叉应用学科,光学、电子学、数学、生物学、计算机技术的迅速发展均对它起到了积极的推动作用.本书在介绍传统的图像处理技术的基础上,以应用实例为背景,引入最新的智能图像处理技术,并对它们的应用方向和发展趋势进行分析和展望.

传统的、工程技术的图像处理方法是建立在模拟和数字信号处理的基础上的,如图像的滤波,用数学上的各种算子对图像进行卷积,或用高通、低通滤波器进行卷积处理,用傅里叶(Fourier)、余弦、子波等正交变换方法,在空间变换域中使信息分布不均匀地集聚在一个小范围中,从而可去除冗余,进行编码,获得高的压缩率;用统计和结构模式识别理论进行图像目标和模式识别等.不管在图像和视频处理上,其理论有多么的完善,但与人的视觉相比,其能力相差很远,如在人的视觉系统中,图像的滤波是用伽博(Gabor)滤波并行进行的.若从能量角度来讨论,人的视觉系统能自动地把能量集中在边缘上,不需要进行数学运算.人的视觉系统每天接收大量的图像和视频信息,视网膜中作为光感受器的是大量的杆体细胞和锥体细胞,约有1.3亿个.光感受器细胞将接收到的光信号转变为神经信号输出,经过神经节细胞、外侧膝状体细胞等的汇集,传送到大脑皮层的视觉区的视觉皮层细胞中.在这一过程中,信息得到大量的压缩,去除了冗余和无用信息,留下真正的精华,为人做各种决策提供依据.这种图像和视频信号的压缩,比计算机做的图像压缩不知要高明多少,其效率也不知要高多少倍;

在图像目标和模式识别的能力上,人的视觉系统更不用说了.因此,智能视觉图像处理方法是模拟人的视觉系统的一种图像处理方法,在工程、军事等领域有着广阔的应用前景.

为了了解智能视觉图像处理方法与传统图像处理方法的区别,在这一章的开始,我们先简单介绍一下传统图像处理方法及其发展,然后在此基础上介绍智能视觉图像处理方法与其不同的地方,并阐述本书的主要思想和思路.通过介绍无监督学习的发展历史与以此引出的以最大熵为基础的独立元分析方法和以亥姆霍兹自由能为基础的单像素独立元分析方法及其独立元分析在生物学上和认知科学上的意义,展示独立元分析方法在国际上的最新研究成果.

第 1 节 传统图像处理方法

人的视觉系统的感知部分——视网膜所感受到的是三维客观世界可见部分的二维投影图像,并根据左右两个二维视图对物体进行三维的理解.对二维图像的三维重建及对三维图像的研究是图像处理领域的一个新的发展热点,但在本章中所介绍的传统图像处理方法只限于二维灰度图像.

在这里,一个静态的灰度图像 I 是指定义在 (x, y) 上的 f 函数的图.

$$I = \{f(x, y), 0 \leq x < L_x, 0 \leq y < L_y\}. \quad (1-1-1)$$

这里, $f: R^2 \rightarrow R$, 而 f 的取值范围为:

$$0 \leq f(x, y) < a. \quad \forall x, y$$

其中, a 是图像灰度的最大值, L_x, L_y 分别是图像的宽和高.对二维空间 x, y 进行抽样(离散化),可获得一个离散化图像,抽样后每个像点的亮度为 $f(i, j), i=1, 2, \dots, N, j=1, 2, \dots, M$, 其中 N, M 分别为图像在 L_x, L_y 上抽样的点数,称为“像素”(pixel).这样一来,二维空间上的连续图像变成了由 $N \times M$ 个像素组成的离散图像,其灰阶也可离散为 k 个等级,灰阶 $f(i, j)$ 变为 $f_n(i, j), 0 \leq$

$f_n(i, j) < k$. 一般情况下, $k = 255$, 每个像素用 8 比特(bit)表示, N, M 可根据不同的分辨率取 $256 \times 256, 325 \times 288, 720 \times 480, 512 \times 512, 1280 \times 720$ 个像素等. 在计算机图像中, 彩色一般采用红(R)、绿(G)、蓝(B)三原色表示(通常为 24 比特). 活动图像的离散化(视频图像)是对 t 时刻采样获得 25 帧/秒、30 帧/秒或 50 帧/秒的视频离散化信号. 图像的采集可通过光学摄像机直接采集, 如 CCD 摄像机、卫星遥感等; 也可用波的反射形成图像, 如雷达、医学 CT 和超声等; 或间接处理获得的图像, 如核磁共振等. 由于图像的获得容易受到光线、人体信号或其他因素的影响, 噪声干扰大, 图像处理一直是人的视觉研究中一个十分重要的领域. 同时, 图像和视频信号的数据量非常大, 可以称为“海量”. 它的信号储存、传输在若干年前始终是图像处理技术得到广泛应用的瓶颈, 因此, 图像压缩方法为众人所关注. 图像目标识别由于在航天、军事、工业中的广泛应用, 也成为图像和视频处理的重要方向. 近几年来, 随着网络技术的不断兴起, 图像传输的加密技术也引起了人们的广泛重视, 逐渐成为许多研究机构的研究热点. 本书介绍的智能视觉图像处理, 囊括了图像处理、图像压缩、图像识别、图像加密等各个不同的领域. 由于在上述领域中引入了智能的概念, 它和传统图像处理方法相比较, 发展前景更为广阔.

1.1 图像处理

图像处理可以定义为将一幅图像变为另一幅经过修改(改进)的图像的过程. 修改(改进)过的图像效果会比原图像好; 或抽取图像的主要特征, 便于识别; 或图像信息更集中, 便于压缩等.

传统图像处理技术基本上可分为: 模拟图像处理和数字图像处理.

模拟图像处理包括: 光学处理(利用透镜)和电子处理, 如照相、遥感图像处理, 电视信号处理等. 模拟图像处理的特点是处理速度快, 并可并行处理, 一般为实时处理, 但缺点是精度较差, 灵活性不强, 其判断能力和非线性处理能力也不够好. 数字图像处理是

对一幅图像的每个像素的灰度的数字表示施加一系列的操作,以得到所希望的结果,一般都用计算机处理或实时硬件处理.其优点是处理精度高,可处理的图像种类多,并可进行复杂的非线性处理,有很强的灵活性.本书主要讨论数字图像处理.

数字图像处理方法可分为两类:空域法和变换域法.空域法主要是在空间域中进行处理;变换域法是图像经过如傅里叶变换、子波变换等进行处理或压缩,主要用于滤波器设计和图像的压缩编码(将在本章第 1.2 节中详细介绍).

空域法是建立在图像和算子卷积的基础上,把图像看成平面中各个像素组成的集合,然后直接对二维函数进行相应的处理的一种方法.空域法主要有以下几类:

点运算、代数运算、几何运算、卷积滤波和二值图像处理等.

1.1.1 点运算

在图像处理中,点运算是简单却很重要的一种空域处理方法,通过改变图像数据占据的灰度范围,使图像变得清晰,尤其在显示图像时,点运算的作用尤为明显.

点运算可以表示为:

$$B(x, y) = f(A(x, y)). \quad (1-1-2)$$

其中 $A(x, y)$ 为输入图像, $B(x, y)$ 为输出图像, f 为灰度变换函数.点运算主要应用于光度矫正(去除图像传感器的非线性影响)、对比度增强(将数字图像中感兴趣的特征从相当窄的灰度范围扩展到可显示灰度的更大范围,这在图像显示中有着非常重要的作用)、图像分割及轮廓线标定(点运算是对于图像进行灰度分割的最为基础的方法,它可以通过直方图分析找出合理的阈值点,并进行划分).

1.1.2 代数运算

代数运算是指对两幅(或两幅以上)输入图像进行点对点的加、减、乘、除等运算,得到输出图像.其数学表达式为:

$$C(x, y) = A(x, y) + B(x, y),$$

$$C(x, y) = A(x, y) - B(x, y),$$

$$C(x, y) = A(x, y) \times B(x, y),$$

$$C(x, y) = A(x, y) \div B(x, y). \quad (1-1-3)$$

其中 $A(x, y)$ 、 $B(x, y)$ 为输入图像, $C(x, y)$ 为输出图像. 代数运算主要应用于图像的简单降噪(图像相加后平均噪声)、二次曝光效果(两幅图像叠加)、运动检测和去除背景(通过减法求得差分图像)、掩膜运算(将掩膜图像和原图像相乘可用于图像分割等)、产生比率图像(除法运算).

1.1.3 几何运算

几何运算是指改变图像中各个物体的几何空间关系, 一般需要两个独立的算法. 首先需要有一个算法定义空间变换, 用来描述每个像素如何移动, 同时还需要用另一个灰度级插值的算法来解决像素映射产生的非整数位置的灰度值计算. 空间变换的算法如下:

$$\begin{aligned} C(x, y) &= f(x', y') = f(a(x, y), b(x, y)), \\ x' &= a(x, y), y' = b(x, y). \end{aligned} \quad (1-1-4)$$

$C(x, y)$ 为输出图像, 函数 $a(x, y)$ 、 $b(x, y)$ 描述了空间变换. 灰度级插值通常采用最近邻插值、双线性插值和高阶插值, 其中双线性插值的应用最为普遍, 如下式所示:

$$C(x, y) = f(ax + by + cxy + d, ex + fy + gxy + h), \quad (1-1-5)$$

即由双线性方程定义的一个双曲抛物面与 4 个已知点拟合. 几何运算常见的应用有几何校正(消除数字图像的几何畸变, 如鱼眼效应)、三维投影的绘制(如卫星图像的处理等)、图像变形(电影、电视中的一些特技处理).

1.1.4 卷积滤波

图像可以看作是一个二维的有限数字信号, 因此, 信号处理的许多技术均可以在数字图像处理中得到应用, 如线性系统理论、信号采样与插值、滤波器的设计等. 其中滤波器的设计最为直观, 也

最为人们所熟悉. 梯度运算 (gradient algorithm)、拉普拉斯算子运算 (Laplace operator)、平滑算子运算 (smoothing operator) 等, 均可用来设计滤波器与图像卷积滤波, 达到去除噪声、增强边缘等目的. 如 $f(x, y)$ 为图像, $h(x, y)$ 为一定窗口的二维滤波算子, $g(u, v)$ 为卷积输出, 其表达式为:

$$g(u, v) = f * h = \int_0^{L_y} \int_0^{L_x} f(x, y) h(u-x, v-y) dx dy, \quad (1-1-6)$$

对于数字化的图像, 其表达式为:

$$g(k, l) = f * g = \sum_{j=0}^{M-1} \sum_{i=0}^{N-1} f(i, j) h(k-i, l-j). \quad (1-1-7)$$

如果 $H(u, v)$ 、 $F(u, v)$ 、 $G(u, v)$ 分别是 $h(x, y)$ 、 $f(x, y)$ 、 $g(x, y)$ 的傅里叶变换, 则在变换域中, 卷积关系变为乘积关系, 即

$$G(u, v) = H(u, v) F(u, v). \quad (1-1-8)$$

若 $f(x, y)$ 是给定的原始数据, 经傅里叶变换后得到 $F(u, v)$, 只要选择合适的 $H(u, v)$, 使由 $g(x, y) = \mathfrak{F}^{-1}[F(u, v) H(u, v)]$ 得到的 $g(x, y)$, 比 $f(x, y)$ 在某些特性方面更加突出, 因而更适合于识别. 这里 $\mathfrak{F}^{-1}$ 为傅里叶反变换.

对于数字图像 f , 设 $N_s(i, j)$ 为 (i, j) 像素的邻近区域,

$$N_s(i, j) = \{k, l / |k-i| \leq r, |l-j| \leq r\}. r \in \{1, 2, 3\}$$

h 的表达式为:

$$h(i, j) = \begin{cases} T(k, l), & k, l \in N_s(i, j), \\ 0, & k, l \notin N_s(i, j). \end{cases} \quad (1-1-9)$$

如以点 (i, j) 为中心, 取 $r=1$ 构成邻域, 则其中所有点的坐标集合为:

$$N_s(i, j) = \{(i, j+1), (i, j-1), (i, j), (i+1, j), (i-1, j), \\ (i-1, j-1), (i+1, j+1), (i-1, j+1), (i+1, j-1)\}.$$

根据不同的处理要求, 可设计不同的 $h(\cdot)$ 对图像进行处理.