



“十三五”科学技术专著丛书

数字图像物体识别 理论详解与实战

柳杨 著

Digital Image Object Recognition:
From Theory to Practice



北京邮电大学出版社
www.buptpress.com



“十三五”科学技术专著丛书

数字图像物体识别 理论详解与实战

柳 杨 著



北京邮电大学出版社
[www. buptpress. com](http://www.buptpress.com)

内 容 简 介

2016年,Alpha Go战胜韩国围棋选手李世乭再次引爆了全球对于人工智能的讨论和关注。计算机视觉作为人工智能技术的基础,受到深度学习的成功影响在近几年内取得了突破性的进展,正在成为影响行业发展的下一个引擎。

巨头们纷纷布局,市场也吸引了越来越多的人才创业参与其中。计算机视觉正在成为人工智能最火热的细分领域之一。

本书详细介绍了最近10年计算机视觉领域中图像识别部分的经典算法和理论。

前半部分分模块,按照预处理、特征提取、特征聚合、分类的顺序介绍了各方向近年来的经典算法。包括图像处理中的去噪、锐化、人脸检测;SIFT、SURF、ORB等局部特征提取算法;全局特征提取算法;BoW、FV、VLAD、稀疏表示等经典特征聚合算法;分类器、贝叶斯理论等模式识别基础知识。理论阐述力求通俗易懂,尽可能扩大读者范围。

后半部分挑选了3种最经典的图像识别算法:HOG+SVM、Sparse Represent Based Classification、Deep Learning;详细讲解了其算法的实现,让读者可以快速实践,更好地学习这些算法。

图书在版编目(CIP)数据

数字图像物体识别理论详解与实战 / 柳杨著. -- 北京:北京邮电大学出版社, 2018.1

ISBN 978-7-5635-5336-5

I. ①数… II. ①柳… III. ①数字图像处理 IV. ①TN911.73

中国版本图书馆CIP数据核字(2017)第310228号

书 名: 数字图像物体识别理论详解与实战

著作责任者: 柳杨 著

责任编辑: 满志文

出版发行: 北京邮电大学出版社

社 址: 北京市海淀区西土城路10号(邮编:100876)

发行部: 电话: 010-62282185 传真: 010-62283578

E-mail: publish@bupt.edu.cn

经 销: 各地新华书店

印 刷: 北京鑫丰华彩印有限公司

开 本: 787 mm×1 092 mm 1/16

印 张: 9.75

字 数: 253千字

版 次: 2018年1月第1版 2018年1月第1次印刷

ISBN 978-7-5635-5336-5

定 价: 38.00元

· 如有印装质量问题,请与北京邮电大学出版社发行部联系 ·

前 言

计算机视觉简介

计算机视觉是一门研究如何使机器“看”的科学,更进一步地说,就是指用摄影机和计算机代替人眼对目标进行识别、跟踪和测量等机器视觉,并进一步做图形处理。作为一个科学学科,计算机视觉研究相关的理论和技术,试图建立能够从图像或者多维数据中获取“信息”的人工智能系统。这里所指的信息指 Shannon 定义的,可以用来帮助做一个“决定”的信息。因为感知可以看作是从感官信号中提取信息,所以计算机视觉也可以看作是研究如何使人工系统从图像或多维数据中“感知”的科学。

计算机视觉是使用计算机及相关设备对生物视觉的一种模拟。它的主要任务就是通过对采集的图片或视频进行处理以获得相应场景的三维信息,就像人类和许多其他类生物每天所做的那样。

计算机视觉是一门关于如何运用照相机和计算机来获取我们所需的,被拍摄对象的数据与信息的学问。形象地说,就是给计算机安装上眼睛(照相机)和大脑(算法),让计算机能够感知环境。我们中国人的成语“眼见为实”和西方人常说的“One picture is worth ten thousand words”表达了视觉对人类的重要性。不难想象,具有视觉的机器的应用前景能有多么的宽广。

计算机视觉既是工程领域,也是科学领域中的一个富有挑战性重要研究领域。计算机视觉是一门综合性的学科,它已经吸引了来自各个学科的研究者参加到对它的研究之中。其中包括计算机科学和工程、信号处理、物理学、应用数学和统计学,神经生理学和认知科学等。

计算机视觉和人工智能

人工智能近几年的火热发展带动了计算机视觉。计算机视觉算法是人工智能最重要的分支之一,它为人工智能提供了眼睛。大部分人工智能依赖计算机视觉算法理解它所处的环境。

未来我们将看到计算机视觉系统与人工智能相结合,来创造并实现各种新的可能性。

这对于我们在计算机视觉领域工作的人来说是一个令人激动的时刻。

希望通过阅读本书,使读者掌握计算机视觉领域的基础知识和经典算法,并能快速实现自己的计算机视觉系统。

目 录

第1章 绪论	1
1.1 人工智能	1
1.2 数字图像处理	2
1.3 数字图像识别	2
1.4 本书架构	3
第2章 图像预处理	4
2.1 引言	4
2.2 数字图像去噪	4
2.2.1 噪声模型	4
2.2.2 利用空间滤波器去噪	7
2.2.3 近几年提出的其他去噪算法	10
2.3 图像锐化	15
2.3.1 图像的微分	15
2.3.2 拉普拉斯算子	16
2.3.3 非锐化掩蔽和高提升滤波	16
2.4 对比度增强	17
2.4.1 直方图均衡	17
2.4.2 指数变换	19
2.4.3 对数变换	20
2.4.4 灰度拉伸	20
2.5 人脸检测	20
2.5.1 简介	21
2.5.2 人脸检测问题分类	21
2.5.3 人脸模式的特征提取	22
2.5.4 人脸模式的特征综合	23
2.5.5 人脸检测总结	27
第3章 特征提取	28
3.1 引言	28
3.2 局部特征提取	28

3.2.1 局部二值模式	28
3.2.2 方向梯度直方图	30
3.2.3 SIFT	33
3.2.4 SURF	40
3.2.5 FAST	42
3.2.6 BRIEF	43
3.2.7 ORB	45
3.3 全局特征提取	45
3.3.1 颜色特征	45
3.3.2 纹理特征	48
3.3.3 形状特征	50
3.3.4 深度学习	54
第4章 特征聚合	62
4.1 引言	62
4.2 词袋模型	62
4.2.1 视觉词袋模型	63
4.2.2 空间金字塔匹配	64
4.3 Fisher Vector	65
4.3.1 Fisher Kernel	65
4.3.2 Fisher Vector	66
4.4 VLAD	67
4.5 合成稀疏模型	68
4.5.1 合成稀疏模型理论	68
4.6 分析稀疏模型	73
4.6.1 分析稀疏模型介绍	74
第5章 模式识别中的分类理论	76
5.1 线性分类模型	76
5.1.1 判别函数	76
5.1.2 概率生成模型	82
5.1.3 概率判别模型	84
5.2 支持向量机	88
5.2.1 简介	88
5.2.2 核方法	92
5.3 神经网络	99
5.3.1 BP 算法	99
5.3.2 梯度消失	100
5.3.3 卷积神经网络	100

第 6 章 实战演练	105
6.1 基于 HOG 的行人检测	105
6.1.1 整体流程	105
6.1.2 实现源码	106
6.1.3 main.cpp	120
6.2 Robust Face Recognition via Sparse Representation	126
6.2.1 核心思想介绍	126
6.2.2 SRC.m	126
6.3 基于深度学习的图像识别	136
6.3.1 LeNet	136
6.3.2 LeNet 的 Caffe 配置文件	138
参考文献	142

第 1 章

绪 论

◆ 1.1 人工智能 ◆

2016 年 3 月阿尔法狗与围棋世界冠军、职业九段选手李世石进行人机大战,并以 4 : 1 的总比分获胜;2016 年年末 2017 年年初,它在中国棋类网站上以“大师”(Master)为注册账号与中日韩数十位围棋高手进行快棋对决,连续 60 局无一败绩。

阿尔法狗的成功引发了人们对人工智能的极大兴趣和热烈的讨论。

那么到底什么是人工智能(Artificial Intelligence, AI)呢? 计算机科学理论奠基人图灵在论文《计算机器和智能》中提出了著名的“图灵测试”——如果一台机器能够与人展开对话(通过电传设备),并且会被人误以为它也是人,那么这台机器就具有智能。人工智能之父之一的马文·明斯基则将其定义为“让机器做本需要人的智能才能够做到的事情的一门科学”。而代表人工智能另一条路线——符号派的司马贺认为,智能是对符号的操作,最原始的符号对应于物理客体。

人工智能其实拥有悠久的发展历史,整个发展史可以概括如下:

- 人工智能的漫长萌芽期(17 世纪~20 世纪 50 年代)。
- 人工智能的诞生(20 世纪 40~50 年代)。
- 人工智能的黄金时代(20 世纪 50~70 年代)。
- 人工智能的第一次低谷(20 世纪 70~80 年代)。
- 人工智能的繁荣期(1980—1987 年)。
- 人工智能的冬天(1987—1993 年)。
- 人工智能的新春(1993 年至今)

随着越来越多的企业开始人工智能的研发,它已经从实验室的学术玩具渗透进日常生活中的方方面面,比如指纹识别、自动车牌识别、人脸支付、AR 红包等。

人工智能中一个重要的分支是计算机视觉^[27],计算机视觉试图让计算机“看懂”图像/视频,指用摄影机和计算机代替人眼完成目标识别、跟踪和测量等一系列任务。而为了让计算机理解图像,数字图像识别是其中的关键步骤。而本书将会从理论和实践的角度对数字图像识别进行详细的介绍。

◆ 1.2 数字图像处理 ◆

图像处理,是对图像进行分析、加工和处理,使其满足视觉、心理以及其他要求的技术。图像处理是信号处理在图像域上的一个应用。目前大多数的图像是以数字形式存储在计算机中的,因而图像处理很多情况下指数字图像处理。

数字图像,是二维图像用有限数字数值像素的表示。数字图像是对图像在空间域和值域上量化的结果。数字图像的基本元素是像素,每一个像素对应着图像上的一个离散位置,而根据每个像素的具体含义,可以将数字图像分为二值图像、灰度图像、彩色图像等。

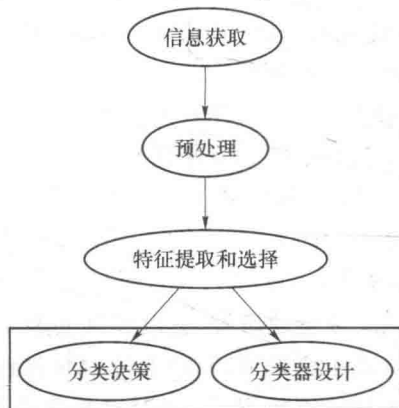
二值图像中每个像素是用 0/1 表示的,通常 0 表示黑,1 表示白。灰度图像的每个像素由一定的比特组成(通常为 8 比特),通过不同的数值,表示该像素点的亮度,一般称为灰度,灰度图像在黑白图像的基础上对每个点的亮度进行细化。彩色图像的每个像素根据不同的颜色空间表示多个通道(常见的为用 24 比特表示 RGB 颜色空间)。

广义的数字图像处理一般分为三个层次:低、中、高。低层次图像处理的特点是处理的输入/输出都是图像,这是狭义的数字图像处理。中层次处理的特点是输入图像,输出特征。高层次处理的特点是模仿人类视觉进行感知。本书将重点关注中和高两个层次的数字图像处理。

◆ 1.3 数字图像识别 ◆

数字图像识别是指利用计算机对数字图像进行处理、分析和理解,以识别各种不同模式的目标和对象的技术^[10]。

数字图像识别的基本流程可以分为五个阶段,如图 1.1 所示。



- (1)信息获取 (2)预处理 (3)特征提取和选择
(4)分类器设计 (5)分类决策

图 1.1 数字图像识别基本流程示意图

信息获取是通过传感器,将光或声音等信息转化为电信息。信息可以是二维的文字、图像等;可以是一维的波形如声波、心电图、脑电图;也可以是物理量与逻辑值^[81]。

预处理包括AD转换、二值化、图像的平滑、变换、增强、恢复、滤波等,主要指狭义数字图像处理或低层次的图像处理。

特征提取和选择是数字图像识别中的关键步骤,它主要负责从数字图像中提取对识别有用的特征,比如颜色特征、纹理特征、形状特征、基于学习的特征等。提取完的特征如果维度过高,为了避免“维度灾难”,一般会进行特征选择,挑选出重要特征以便进行后续流程。^[9]

分类器设计的主要功能是通过训练确定判决规则,使按此类判决规则分类时,错误率最低。

分类决策是在特征空间中对被识别对象进行分类。

◆ 1.4 本书架构 ◆

本书将从理论和实践两个角度对数字图像处理这个主题进行阐述。

全书共分六章:

- 第1章 绪论^[84]。以人工智能为切入点引入数字图像处理的基本概念,并对全书进行介绍。
- 第2章 图像预处理。详细说明了图像去噪、图像锐化、对比度增强、轮廓提取等图像预处理的理论。
- 第3章 特征提取。详细阐述了局部特征提取和全部特征提取两大类方法。
- 第4章 特征聚合。如果选择局部特征提取,则每张图像的局部特征数量未知,需要通过特征聚合的方法得到每一张图片的定长特征向量。第4章详细介绍了目前主流学术论文中的特征聚合理论。
- 第5章 分类器。本章由简入繁,介绍了 k NN、逻辑回归、支持向量机、人工神经网络等分类器的推导、意义,训练方法等。
- 第6章 数字图像识别实战。本章充分利用前面所阐述的理论知识,以行人检测、人脸识别、基于深度学习的图像识别这3个项目为实例,完成数字图像识别的实践。

第 2 章

图像预处理

◆ 2.1 引言 ◆

图像预处理一般都是底层次的图像处理,也就是输入是数字图像,输出也是数字图像。预处理的目的是更好地完成后续的任务。比如在被噪声污染的图像上进行的去噪操作、为了获得更清晰的图像而进行的锐化操作、为了弥补照片曝光时间不够而进行的对比度增强操作等。更进一步地,在预处理过程中,也有一些中层次的图像处理操作,比如在人脸识别之前要进行的人脸检测、在图像分割和图像识别中可能会用到的轮廓提取等。

◆ 2.2 数字图像去噪 ◆

2.2.1 噪声模型

数字图像中,噪声最主要是来自于图像的获取以及图像的传输。当进行成像时,成像传感器很有可能受到外界各种因素的干扰,比如图像获取过程中的环境条件变化,在使用 CCD 摄像机获取图像时,光照水平和传感器温度是影响结果图像中噪声数量的主要因素。图像在传输中被污染主要是由于传输信道中的干扰,例如,使用无线网络传输的图像可能会因为光照或其他大气因素而被污染。

假设干净的图像信号是 $f(x, y)$, 那观测和传输的过程可以简单地用下式描述:

$$g(x, y) = f(x, y) + n(x, y) \quad (2.1)$$

式中, $n(x, y)$ 是加性噪声, $g(x, y)$ 是观测到的图像。

一般来说无法准确地给出噪声 $n(x, y)$ 的性质,一般都将噪声建模为随机变量。

在数字图像处理中,常见的噪声分布如下。

1. 高斯噪声

高斯噪声在数学上容易处理,中心极限定理又告诉我们“大量随机变量之和近似服从正态分布”,因此在实践中高斯噪声(又称为正态噪声)是最常用的。

高斯噪声的概率密度函数如下:

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-\bar{z})^2/2\sigma^2} \quad (2.2)$$

式中, z 表示灰度, σ 是高斯分布的标准差。

高斯分布的概率密度函数如图 2.1 所示。

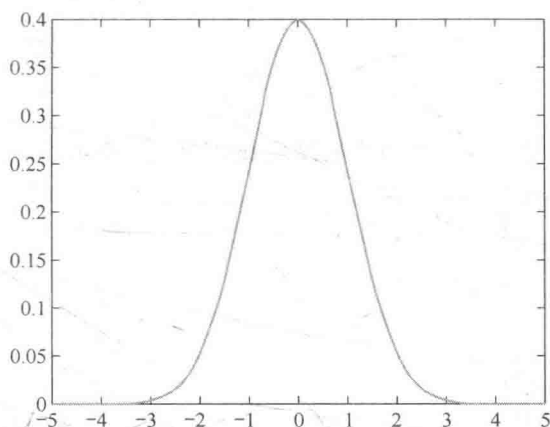


图 2.1 高斯分布概率密度函数图示

2. 瑞利噪声

瑞利噪声的概率密度函数如下:

$$p(z) = \begin{cases} \frac{2}{b}(z-a)e^{-(z-a)^2/b}, & z \leq a \\ 0, & z < a \end{cases} \quad (2.3)$$

瑞利噪声的均值和方差为

$$\bar{z} = a + \sqrt{\pi b/4} \quad (2.4)$$

$$\sigma^2 = \frac{b(4-\pi)}{4} \quad (2.5)$$

当 $a=1, b=2$ 时, 瑞利分布的概率密度函数如图 2.2 所示。

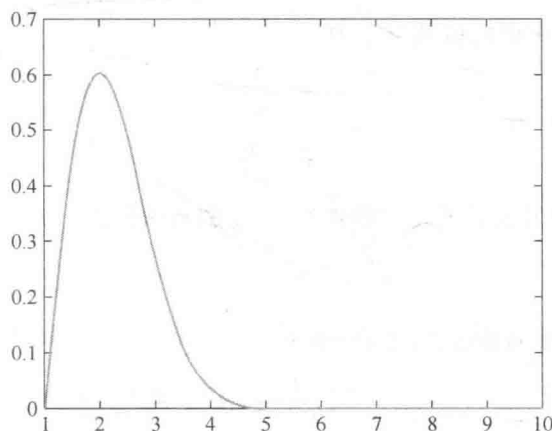


图 2.2 瑞利分布概率密度函数图示

3. 爱尔兰(伽马)噪声

爱尔兰噪声的概率密度函数如下:

$$p(z) = \begin{cases} \frac{a^b z^{b-1}}{(b-1)!} e^{-az}, & z \leq a \\ 0, & z > a \end{cases} \quad (2.6)$$

式中参数 $a > 0, b$ 为正整数, $!$ 表示阶乘, 其均值和方差为^[68]:

$$\bar{z} = \frac{b}{a} \quad (2.7)$$

$$\sigma^2 = \frac{b^2}{a} \quad (2.8)$$

当 $a=2, b=7$ 时, 爱尔兰噪声的概率密度函数如图 2.3 所示。

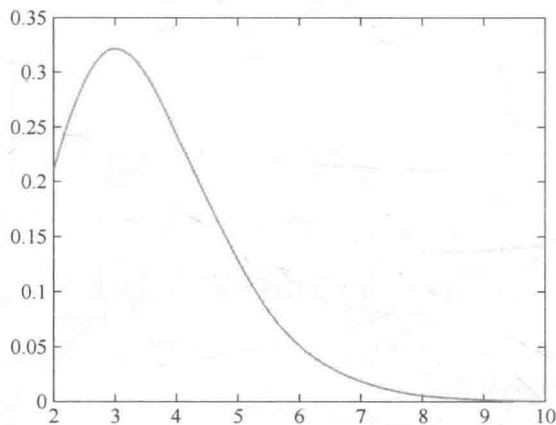


图 2.3 爱尔兰分布概率密度函数图示

尽管式(2.6)经常被称为伽马密度, 但是严格地说, 这只有在分母为伽马函数 $\Gamma(b)$ 时才是正确的。式(2.6)被称为爱尔兰密度更合适。

4. 指数噪声

指数噪声的概率密度函数由式(2.9)给出:

$$p(z) = \begin{cases} a e^{-az}, & z \leq 0 \\ 0, & z > 0 \end{cases} \quad (2.9)$$

式中 $a > 0$ 。指数噪声的均值和方差为

$$\bar{z} = \frac{1}{a} \quad (2.10)$$

$$\sigma^2 = \frac{1}{a^2} \quad (2.11)$$

可以看出, 指数噪声其实是爱尔兰噪声在 $b=1$ 的特殊情况。 $a=1$ 时, 指数噪声的概率密度函数如图 2.4 所示。

5. 均匀噪声

均匀噪声的概率密度函数由式(2.12)给出:

$$p(z) = \begin{cases} \frac{1}{b-a}, & a \leq z \leq b \\ 0, & \text{其他} \end{cases} \quad (2.12)$$

其中 $a > 0$ 。均匀噪声的均值和方差为

$$\bar{z} = \frac{a+b}{2} \quad (2.13)$$

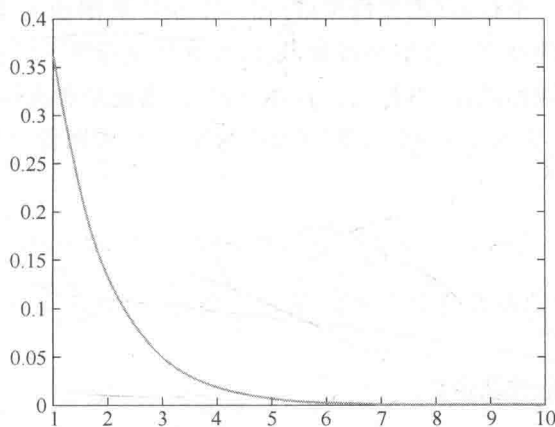


图 2.4 指数分布概率密度函数图示

$$\sigma^2 = \frac{(b-a)^2}{12} \quad (2.14)$$

$a=1, b=7$ 时, 均匀噪声的概率密度函数如图 2.5 所示。

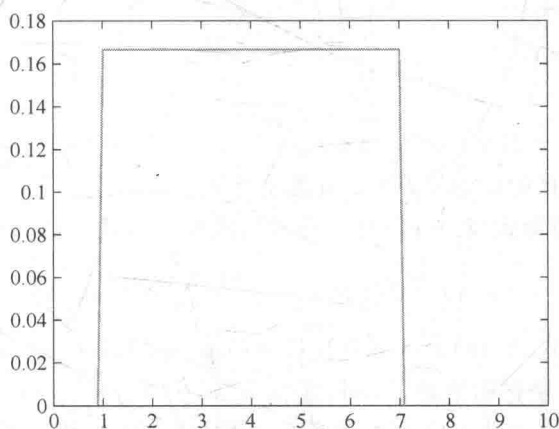


图 2.5 均匀分布概率密度函数图示

6. 椒盐噪声

椒盐噪声的概率密度函数如下:

$$p(z) = \begin{cases} P_a, & z=a \\ P_b, & z=b \\ 1-P_a-P_b, & \text{其他} \end{cases} \quad (2.15)$$

如果 $b > a$, 则 b 表示一个亮点, a 表示一个暗点。当 P_a 和 P_b 近似相等, 且不为 0 时, 该噪声类似随机分布在图像上的胡椒和盐粉, 这也是椒盐噪声这个名字的由来。

2.2.2 利用空间滤波器去噪

1. 空间滤波

空间滤波是数字图像处理领域应用广泛的工具之一, 本小节首先对空间滤波进行简单介绍^[51]。

空间滤波操作是用一个模板(或者称为核)在一小块矩形图像中进行运算,最后输出一个像素的操作。模板或核的宽高一般都是奇数,输出像素的坐标等于模板中心位置的坐标,如果运算中对像素执行的是线性操作,则称之为线性滤波,否则就是非线性滤波。

一般来说,使用大小为 $m \times n$ 的滤波器对大小为 $M \times N$ 的图像进行线性空间滤波的公式如下:

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x+s, y+t) \quad (2.16)$$

式中, x 和 y 就对应滤波器的中心位置,它是可变的^[77],以便 w 中的每个像素可以访问 f 中的每个像素。

比如使用一个如下的滤波器:

$$\begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \quad (2.17)$$

对如下矩阵做滤波:

$$\begin{pmatrix} 8 & 3 & 1 \\ 3 & 5 & 6 \\ 2 & 7 & 4 \end{pmatrix}$$

得到的结果为

$$8 \times (-1) + 3 \times (-2) + 1 \times (-1) + 3 \times 0 + 5 \times 0 + 6 \times 0 + 2 \times 1 + 7 \times 2 + 4 \times 1 = 5$$

空间滤波和卷积运算的概念很像^[92],需要注意区分^[2]。

为了方便比较,采用和式(2.16)同样的符号写出卷积运算的公式如下:

$$g(x, y) = w(x, y) \times f(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, y) f(x-s, y-t) \quad (2.18)$$

注意到式(2.18)和式(2.16)的区别在最后一项上,滤波是 $f(x+s, y+t)$ 而卷积是 $f(x-s, y-t)$ 。更直观的解释是它们俩差了一个镜像: $w(x, y)$ 在 $f(x, y)$ 上做滤波相当于在 $f(-x, -y)$ 上做卷积。不过很多情况下空间滤波器是对称的,所以卷积和滤波的结果一样,这就导致很多人以为卷积和滤波是同样的运算^[94]。

2. 滤波去噪

滤波是一种很基本的去噪方式,但是为什么通过滤波可以去噪呢?本节将给出一个简单的解释。

首先,不是所有的滤波器都可以用来去噪,前一小节中的式(2.17)一般就不能用来去噪,它的作用主要是边缘提取,这类滤波器从频域的角度来看,也被称为高通滤波器。

一般可以用来去噪的都是模糊滤波器,或者被称之为低通滤波器。典型的例子就是均匀滤波器,下面给出一个 3×3 的均匀滤波器:

$$\frac{1}{9} \times \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (2.19)$$

而且模糊滤波器也不是对所有图像去噪效果都很好,它一般是在图像的平坦区域具有较好的去噪效果,下面给出简单的理论解释。

假设干净的图像为 f ,加性噪声为 n ,服从零均值的高斯分布,观测得到的图像为 $g=f+$

n 。那通过 3×3 的均匀滤波器滤波之后得到结果: $g' = f' + n'$ 。假设 f 为平坦的图像区域, 则模糊滤波器对它的影响可以忽略, 或者说: $f' = f$, 所以 $g' \approx f + n'$ 。

而 $n' = \sum_{i=0}^8 n_i$, 是其邻域点加上自身的平均。我们知道 n 是服从零均值高斯分布, 那 n' 是独立同分布高斯分布的和, 它也是零均值, 但是 n' 的方差由原来的 σ^2 变成了 $\frac{\sigma^2}{9 \times 9}$ [66]。

噪声的方差变小了, 所以模糊滤波器对图像的平坦区域具有较好的去噪效果。

另外一方面, 可以从频域的角度去解释。一般来说, 都假设噪声是高斯白噪声, 白就意味着其拥有所有的频率分量。而模糊滤波器(也就是低通滤波器)将噪声的高频部分滤掉, 从而去掉了部分噪声。

虽然低通滤波器可以一定程度上的去噪, 但是它只能去掉“部分”噪声, 而且只在图像“平坦”的区域有较好的效果。

图 2.6 给出了使用高斯低通滤波器去除噪声的效果示意图。

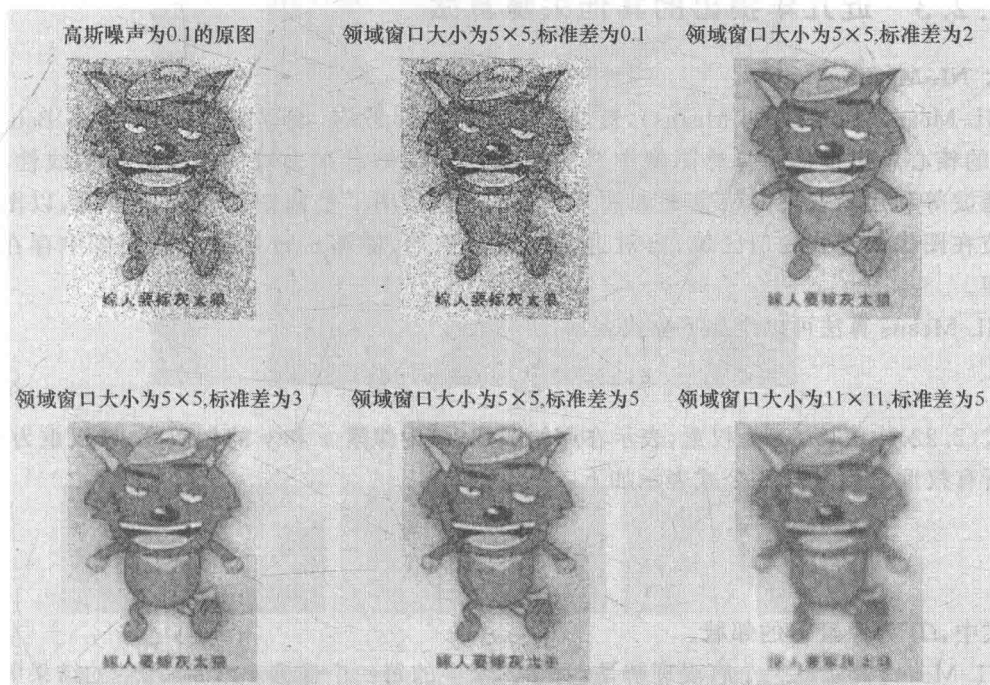


图 2.6 高斯低通滤波器去噪的效果示意图

3. 利用空间滤波器去噪

上一节简单解释了空间滤波器的去噪原理, 并给出了高斯滤波器去噪的例子。

本节将介绍其他的空间滤波器的例子。

几何平均滤波器。式(2.19)给出了 3×3 均匀滤波器的例子, 不过更严格地说, 这种滤波器称为算术均值滤波器, 因为它求得是算术平均。与之对应的还有几何平均, 几何平均滤波器就是一种求图像块的几何平均数并得到结果的滤波器, 其滤波器表示式如下:

$$f(x, y) = \left[\prod_{(s, t) \in S(x, y)} g(s, t) \right]^{\frac{1}{mn}} \quad (2.20)$$

式中, m 和 n 表示图像块的宽高, $S(x, y)$ 表示该图像块。和算术平均滤波器相比, 几何平均滤波器丢失的图像细节更少。

中值滤波器。除了算术平均和几何平均外,在统计学中,还有一个中值的概念,是指对一个序列进行排序,在排序后的结果中,取位置在中间的值。中值滤波器比同尺寸的模糊滤波器的模糊效应要小一些,而且它对椒盐噪声的效果非常显著。

自适应滤波器。滤波器作用于局部区域 S_{xy} 。滤波器在该区域中心任意一点 (x, y) 上的响应基于以下 4 个量:① $g(x, y)$, 带噪图像在点 (x, y) 上的值;② σ_g^2 , 该领域内噪声的方差;③ m_L , S_{xy} 中像素的局部均值;④ σ_L^2 , S_{xy} 中像素的局部方差。而希望的滤波器性能如下:

- 如果 σ_g^2 为 0, 则滤波器应该返回 $g(x, y)$ 。
- 如果局部方差和 σ_g^2 高度相关, 则滤波器应该返回 $g(x, y)$ 的一个相似值。
- 如果两个方差相等, 则滤波器应该返回 S_{xy} 中像素的算术平均值。

基于以上假设,可以得到自适应滤波器的公式如下:

$$f(x, y) = g(x, y) - \frac{\sigma_g^2}{\sigma_L^2} [g(x, y) - m_L] \quad (2.21)$$

2.2.3 近几年提出的其他去噪算法

1. NL-Means

NL-Means(Non-Local Means),直译过来是非局部平均,该算法在 2005 年由 Baude 提出,它的核心思想是使用自然图像中普遍存在的冗余信息来去噪声。与常用的线性滤波、中值滤波等利用图像局部信息来滤波不同的是,它利用了整幅图像来进行去噪,以图像块为单位在图像中寻找相似区域,再对这些区域求平均,能够比较好地去掉图像中存在的高斯噪声。

NL-Means 算法可以用如下公式表示:

$$\tilde{u}(x) = \sum_{y \in \Omega_x} w(x, y) v(y) \quad (2.22)$$

式(2.22)中 $w(x, y)$ 为权重,表示在原始图像 (v) 中像素 x 和 y 的相似度,该权重为正数,并且所有权重之和为 1,用公式表示如下:

$$\begin{aligned} w(x, y) &> 0 \\ \sum_{y \in \Omega_x} w(x, y) &= 1 \end{aligned}$$

式中, Ω_x 为像素 x 的邻域。

NL-Means 算法的一个直观理解是:对于图像中的每一个像素 x ,去噪之后的结果等于它邻域中像素 y 的加权和,加权的权重等于 x 和 y 的相似度。这个邻域也称为搜索区域,搜索区域越大,找到相似像素的机会也越大,但同时计算量也是成指数上升。

从上面的描述可以看出,NL-Means 算法中的一个关键点就是如何计算权重 $w(x, y)$,或者说如何衡量两个像素的相似度。衡量像素相似度的方法有很多,最常用的是根据两个像素的亮度值的差的平方来估计。但因为有噪声的存在,单独的一个像素并不可靠。对此解决方法是,考虑它们的邻域,只有邻域相似度高才能说这两个像素的相似度高。衡量两个图像块的相似度最常用的方法是计算他们之间的高斯加权欧氏距离,具体公式如下:

$$w(x, y) = \exp\left(-\frac{\|P(x) - P(y)\|_{2a}^2}{\sigma^2}\right) \quad (2.23)$$

式中, σ 是高斯函数中的方差, $P(x)$ 和 $P(y)$ 代表了像素 x 和像素 y 的邻域,这个邻域常称为块邻域,块邻域一般要小于搜索区域。 $\|P(x) - P(y)\|_{2a}^2$ 是以 x 和以 y 为中心的两个块中各