

提供配套实验仿真
源代码下载

视觉机器学习

20讲

谢剑斌 等编著

视觉机器学习算法理论与实践有机结合，解决视觉机器学习领域中的诸多基础问题

重视算法的典型性和可实现性，既包含本领域的经典算法，也包含本领域的最新研究成果



清华大学出版社

视觉机器学习

20讲

谢剑斌 兴军亮 张立宁 方宇强 李沛秦 刘通 编著
闫玮 王勇 沈杰 张政 潘筠 胡俊



清华大学出版社
北京

内 容 简 介

本书是计算机、自动化、信息、电子与通信学科方向的专著,详尽地介绍了K-Means、KNN学习、回归学习、决策树学习、Random Forest、贝叶斯学习、EM算法、Adaboost、SVM方法、增强学习、流形学习、RBF学习、稀疏表示、字典学习、BP学习、CNN学习、RBM学习、深度学习、遗传算法、蚁群方法等基本理论;深入阐述了视觉机器学习算法的优化方法和实验仿真;系统地总结了其优点和不足。

本书特别重视如何将视觉机器学习算法的理论和实践有机地结合,解决视觉机器学习领域中的诸多基础问题,可应用于医学图像分析、工业自动化、机器人、无人车、人脸检测与识别、车辆信息识别、行为检测与识别、智能视频监控等。本书特别重视算法的典型性和可实现性,既包含本领域的经典算法,也包含本领域的最新研究成果。

本书不仅可作为高年级本科生与研究生教材,而且也是从事视觉机器学习领域研发极为有用的参考资料。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

视觉机器学习 20讲 / 谢剑斌等编著. —北京:清华大学出版社, 2015

ISBN 978-7-302-39793-3

I. ①视... II. ①谢... III. ①机器学习 IV. ①TP181

中国版本图书馆CIP数据核字(2015)第080934号

责任编辑:王金柱

封面设计:王翔

责任校对:闫秀华

责任印制:王静怡

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者:北京富博印刷有限公司

装 订 者:北京市密云县京文制本装订厂

经 销:全国新华书店

开 本:180mm×230mm

印 张:16.25

字 数:364千字

版 次:2015年6月第1版

印 次:2015年6月第1次印刷

印 数:1~3000

定 价:49.00元

本书作者简介



谢剑斌，博士，教授，国防科学技术大学电子科学与工程学院研究生导师，中国生物特征识别国家标准组委员，中国数字电视国家标准组委员，中国图像图形学会高级会员，计算机学会高级会员，电光与控制编委会委员。长期从事海量视频分析与生物特征识别方面的研究工作，作为项目负责人主持国家级项目 8 项、部委级项目 27 项、横向课题 39 项。在国内外知名期刊发表学术论文 100 多篇，出版专著 6 部，授权国家发明专利 21 项、实用新型专利 32 项。荣获国际“发明展”金奖 2 项、湖南省科技进步二等奖 1 项、公安部技术革新特别项目奖 2 项、全国“发明展”金奖 2 项、铜奖 1 项。



兴军亮，2012 年毕业于清华大学计算机科学与技术系，获工学博士学位，2013 年在新加坡国立大学从事博士后研究，现为中国科学院自动化研究所助理研究员，美国电器与电子工程学会（IEEE）会员。研究领域为计算机视觉、模式识别和机器学习，主要研究兴趣为视频中的物体检测、跟踪和分割。目前已在重要国际期刊和会议，如 TIP, ICCV, CVPR, ACM Multimedia 上发表论文多篇，译著两部。在攻读博士期间，曾获得清华大学计算机系学术新秀、清华大学优秀博士论文、清华大学综合一等奖学金等荣誉称号和奖励，在博士后研究期间，带领学生完成的论文 *Wow! You are so beautiful today!* 获得 2013 年 ACM Multimedia 国际会议最佳论文奖。目前作为项目负责人正在主持一项国家自然科学基金项目，作为项目骨干参与了 973、863 等多项国家重要研究课题以及与惠普、英特尔、欧姆龙等知名企业的多项合作研究课题。



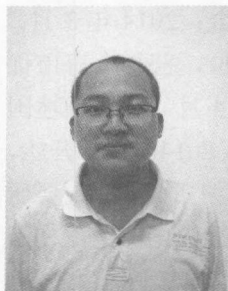
张立宁, 2006 年和 2009 年毕业于西安电子科技大学并分别获学士和硕士学位, 2013 年毕业于新加坡南洋理工大学电气与电子工程学院, 获博士学位, 现为新加坡科技研究局资讯通信研究院研究员, 美国电气与电子工程师协会会员。研究领域为模式识别与机器视觉, 多媒体计算, 医学图像处理等。目前已在国际重要刊物如 IEEE TIP、TCSVT、TSMC-PartB 等发表论文多篇, 参与编辑著作多部。在攻读硕士和博士学位期间, 曾获多项校优秀学生、优秀毕业生等荣誉称号, 并作为骨干成员参与中国 973、863 多项重要研究课题以及新加坡多媒体创新的研究项目。



方宇强, 2010 年毕业于国防科学技术大学控制科学与工程专业, 获得硕士学位, 现就读于国防科学技术大学控制科学与工程专业, 攻读博士学位。2011 年赴香港科技大学计算机科学系访问, 2013 年在新加坡国立大学计算机系任研究助理。研究领域包括机器学习、计算机视觉和智能系统, 主要研究兴趣有高维数据特征分析、视觉特征学习和智能无人系统等。在国内外重要期刊和会议, 如 ICDM、JFR 上发表多篇论文, 硕士论文获得湖南省优秀硕士毕业论文。在攻读博士期间作为项目骨干参与了多项国家 973、国家自然科学基金重点项目研究。



李沛秦, 博士, 讲师, 任职于国防科学技术大学电子科学与工程学院。长期从事海量视频分析与生物特征识别方面的研究工作, 作为技术骨干参与国家级项目 3 项, 省部级项目 13 项。在国内外知名期刊发表学术论文 25 篇, 参与出版专著 2 部, 授权国家发明专利 12 项、实用新型专利 25 项。荣获国际发明展金奖 2 项、湖南省科技进步二等奖 1 项、公安部技术革新特别项目奖 2 项、全国“发明展”金奖 1 项、铜奖 1 项。



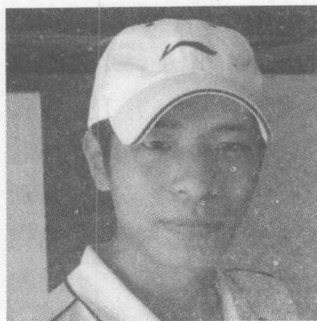
刘通, 博士, 讲师, 任职于国防科学技术大学电子科学与工程学院。长期从事海量视频分析与生物特征识别方面的研究工作, 作为技术骨干参与国家级项目 3 项, 省部级项目 17 项。在国内外知名期刊发表学术论文 29 篇, 参与出版专著 2 部, 授权国家发明专利 12 项、实用新型专利 25 项。荣获国际“发明展”金奖 2 项、湖南省科技进步二等奖 1 项、公安部技术革新特别项目奖 2 项、全国“发明展”金奖 1 项、铜奖 1 项。



闫玮, 博士, 讲师, 任职于国防科学技术大学电子科学与工程学院。长期从事海量视频分析与生物特征识别方面的研究工作, 作为技术骨干参与国家级项目 3 项, 省部级项目 15 项。在国内外知名期刊发表学术论文 23 篇, 参与出版专著 2 部, 授权国家发明专利 12 项、实用新型专利 25 项。荣获国际“发明展”金奖 2 项、湖南省科技进步二等奖 1 项、公安部技术革新特别项目奖 2 项、全国“发明展”金奖 1 项、铜奖 1 项。



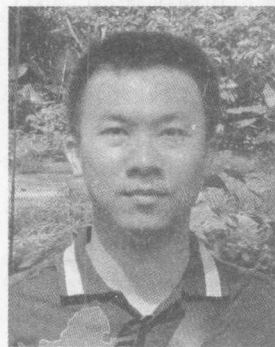
王勇, 1982 年 2 月生。2005 年 7 月于四川大学数学学院获学士学位, 分别于 2007 年 12 月和 2011 年 6 月于国防科学技术大学理学院获硕士和博士学位。2011 年 7 月起在中国空气动力研究与发展中心低速所任助理研究员。研究兴趣包括机器学习、统计学习、模式识别与计算机视觉, 论文 *Multi-manifold clustering* 于 2010 年获亚太人工智能大会最佳论文奖 (Best Paper Award), 论文 *Face recognition using Intrinsicfaces* 被国际期刊 *Pattern Recognition* 评选为该刊 2010 年第二季度最热门的 25 篇文章之一 (Top 25 Hottest Articles), 排行第 14。目前主要从事气动声学、机器学习等领域的研究工作。



沈杰, 博士生, 于 2011 年 6 月在上海交通大学数学系获得理学学士学位, 多次获得学业优秀奖学金; 2014 年 3 月, 在上海交通大学计算机系获得工学硕士学位, 获得上海市优秀毕业生称号; 从 2013 年 8 月至 2014 年 4 月, 在新加坡国立大学视觉和机器学习实验室访问研究; 从 2014 年 4 月开始, 在 USA Rutgers 大学师从 Ping Li 教授攻读博士学位。主要研究人物姿势识别、物体识别、特征学习和凸优化理论, 在凸优化理论的框架下, 设计简单高效的算法。



张政, 1983 年 2 月生。2005 年 7 月于空军第一航空学院获学士学位, 于 2007 年 12 月获国防科学技术大学航天科学与工程学院获硕士学位, 于 2012 年 12 月获南洋理工 (Nanyang Technological University) 计算机工程学院博士学位。2013 年 1 月起在国防科学技术大学信息系统与管理学院任讲师。在 IEEE Transactions on Multimedia、ICIP 等国际期刊和会议上以第一作者发表论文近 10 篇。研究兴趣包括计算机视觉、图像处理、机器学习等。目前主要从事基于视觉的智能监控、巨量视频数据处理和分析等领域的研究工作。



谭筠, 博士, 2007 年毕业于国防科技大学机电工程与自动化学院, 获工学学士学位。2010 年暑期, 在第一汽车集团公司短期交流学习, 合作研制改进了新一代的导航辅助系统。2011 年 11 月至 2013 年 2 月获中国国家留学基金委资助, 在新加坡国立大学视觉研究组访问学习。研究领域为计算机视觉、模式识别和机器学习, 主要研究无人驾驶车辆、汽车主动安全产品中的视觉感知问题。作为主要成员, 获得了 2011 中国智能车未来挑战赛第一名, ICPR2012 人类行为识别与定位大赛第一名等国内、国际奖项。在国际、国内主流会议上发表多篇学术论文, 获得第十届中国智能机器人学术会议的优秀论文奖。作为项目骨干, 参加了国家自然科学基金重大项目、973、863 等多项国家重要研究课题。参与开发一系列的汽车主动安全产品, 部分已经产业化。



胡俊, 1985 年 6 月生。2008 年 6 月于湖南大学电气与信息工程学院获学士学位, 2011 年 12 月于国防科学技术大学电子科学与工程学院获硕士学位, 并于 2012 年 2 月起在国防科学技术大学电子科学与工程学院攻读博士学位。研究兴趣包括信息信号处理、目标检测、识别与跟踪。

前言

在新加坡国立大学访问期间，我经常和兴军亮博士、张立宁博士、方宇强博士、谭筠博士等讨论，谈到视觉机器学习非常有用，但是很难找到称心如意的参考资料，网络上相关资源不少，但是有的泛泛而谈，有的空洞无味，有的实例无法运行，有的效率很低，有的缺乏实验数据……曾经甚是苦闷，回过头来思考，大家或多或少都有些酸甜苦辣的经验和教训。我提议大家就常用视觉机器学习方法，贡献出自己的收获，获得大家一致赞同。于是我们在多次会议讨论之后，形成视觉机器学习 20 讲的提纲，大家分头撰写，然后由我来统稿、修改和完善。经过近一年的努力，终于开花结果，真心希望这本读书笔记和心得体会能够让同道中人少些奔走，免除去粗取精、去伪存真的劳苦，熟悉算法的来源、发展和所以然，掌握算法的改进方法、实验仿真流程、源代码和视频库，使得研究人员和学生们尽快上手，树立深入研究的信心。

本书是视觉机器学习领域的 12 位博士多年研究视觉机器学习的基础理论、核心算法、关键技术和应用实践的科学结晶，是作者们多年相关科研实践的心得体会和系统总结。包括 K-Means、KNN 学习、回归学习、决策树学习、Random Forest、贝叶斯学习、EM 方法、AdaBoost、SVM 方法、增强学习、流形学习、RBF 学习、稀疏表示、字典学习、BP 学习、CNN 学习、RBM 学习、深度学习、遗传方法、粒子群方法、蚁群方法共 20 个常用视觉机器学习方法。人们对其进行了深入浅出的阐述，以简单明了、丰富图表、解释代码、面向应用、服务研究生等为宗旨，从基本原理、实现方法、改进方法、仿真流程、核心代码、源程序、实验数据等方面重点展开，适合于从事医学图像分析、工业自动化、机器人、无人车、人脸检测与识别、车辆信息识别、行为检测与识别、智能视频监控等领域的研究生和技术人员学习参考。

本书章节安排

全书共 1 个绪论，20 讲内容，各章内容编写安排如下：

绪论 (谢剑斌博士);

第 1 讲 K-Means (方宇强博士);

第 2 讲 KNN 学习 (谢剑斌博士);

第 3 讲 回归学习 (闫玮博士);

第 4 讲 决策树学习 (谢剑斌博士);

第 5 讲 Random Forest (张政博士);

第 6 讲 贝叶斯学习 (刘通博士);

第 7 讲 EM 算法 (王勇博士);

第 8 讲 Adaboost (胡俊博士);

第 9 讲 SVM 方法 (张立宁博士);

第 10 讲 增强学习 (李沛秦博士);

第 11 讲 流形学习 (张立宁博士);

第 12 讲 RBF 学习 (谢剑斌博士);

第 13 讲 稀疏表示 (方宇强博士);

第 14 讲 字典学习 (方宇强博士);

第 15 讲 BP 学习 (兴军亮博士);

第 16 讲 CNN 学习 (兴军亮博士);

第 17 讲 RBM 学习 (兴军亮博士);

第 18 讲 深度学习 (兴军亮博士);

第 19 讲 遗传算法 (张立宁博士);

第 20 讲 蚁群方法 (谢剑斌博士);

沈杰博士、谭筠博士在实验仿真方面做了很多工作。

本书还提供了实验仿真源代码资源下载, 下载地址为 www.kedachang.com 或 <http://pan.baidu.com/s/1c051viS>。

本书由国防科技大学电子科学与工程学院数字视频课题组组稿, 谢剑斌教授负责全书审核。在编著过程中得到了国防科技大学庄钊文教授、唐朝京教授的大力支持, 穆春迪、刘双亚、李润华等为本书编著做了大量工作, 在此一并致谢! 由于时间有限, 参与作者较多, 可能没有列全参考文献, 请读者或相关作者来信告知, 在提供配套资源的网站和再版时加入并致谢。

编者

目录

绪论	1
第 1 讲 K-means	11
1.1 基本原理	11
1.2 算法改进	13
1.3 仿真实验	16
1.4 算法特点	18
第 2 讲 KNN 学习	20
2.1 基本原理	20
2.2 算法改进	23
2.3 仿真实验	24
2.4 算法特点	26
第 3 讲 回归学习	28
3.1 基本原理	28
3.1.1 参数回归	29
3.1.2 非参数回归	30
3.1.3 半参数回归	30
3.2 算法改进	30

3.2.1	线性回归模型	30
3.2.2	多项式回归模型	31
3.2.3	主成分回归模型	32
3.2.4	自回归模型	33
3.2.5	核回归模型	33
3.3	仿真实验	37
3.3.1	回归学习流程	37
3.3.2	基于回归学习的直线边缘提取	37
3.3.3	基于回归学习的图像插值	39
3.4	算法特点	41
第 4 讲	决策树学习	42
4.1	基本原理	42
4.1.1	分类与聚类	42
4.1.2	决策树	43
4.1.3	信息增益的度量标准	43
4.1.4	信息增益度量期望的熵降低	44
4.1.5	悲观错误剪枝 PEP	46
4.1.6	基本决策树算法	47
4.2	算法改进	47
4.2.1	ID3 算法	47
4.2.2	C4.5 算法	48
4.2.3	SLIQ 算法	49
4.2.4	SPRINT 算法	49
4.3	仿真实验	50
4.3.1	用于学习布尔函数的 ID3 算法伪代码	50
4.3.2	C4.5 算法构造决策树的伪代码	51
4.4	算法特点	53
第 5 讲	Random Forest 学习	54
5.1	基本原理	54
5.1.1	决策树	55

5.1.2	Bagging 集成学习	55
5.1.3	Random Forest 方法	56
5.2	算法改进	57
5.3	仿真实验	58
5.3.1	Random Forest 分类与回归流程	58
5.3.2	Forest-RI 和 Forest-RC	59
5.3.3	基于 Random Forest 的头部姿态估计	59
5.4	算法特点	60
第 6 讲	贝叶斯学习	62
6.1	基本原理	62
6.2	算法改进	63
6.2.1	朴素贝叶斯模型	63
6.2.2	层级贝叶斯模型	65
6.2.3	增广贝叶斯学习模型	66
6.2.4	基于 Boosting 技术的朴素贝叶斯模型	66
6.2.5	贝叶斯神经网络模型	66
6.3	仿真实验	66
6.3.1	Learn_Bayse(X,V)	67
6.3.2	Classify_Bayse(X)	67
6.4	算法特点	68
第 7 讲	EM 算法	70
7.1	基本原理	70
7.2	算法改进	71
7.2.1	EM 算法的快速计算	71
7.2.2	未知分布函数 $P_i(y_i)$ 的选取	74
7.2.3	EM 算法收敛性的改进	75
7.3	仿真实验	77
7.3.1	EM 算法流程	77
7.3.2	EM 算法的伪代码	77
7.3.3	EM 算法应用——高斯混合模型	77

7.4 算法特点	79
第 8 讲 Adaboost	81
8.1 基本原理	81
8.1.1 Boosting 方法	81
8.1.2 Adaboost 方法	82
8.2 算法改进	83
8.2.1 权值更新方法的改进	83
8.2.2 Adaboost 并行算法	83
8.3 仿真实验	83
8.3.1 Adaboost 算法实现流程	83
8.3.2 Adaboost 算法示例	84
8.4 算法特点	86
8.4.1 Adaboost 算法的优点	86
8.4.2 Adaboost 算法的缺点	87
第 9 讲 SVM 方法	88
9.1 基本原理	88
9.2 算法改进	90
9.3 仿真实验	94
9.4 算法特点	100
第 10 讲 增强学习	102
10.1 基本原理	102
10.2 算法改进	105
10.2.1 部分感知模型	105
10.2.2 增强学习中的函数估计	105
10.2.3 分层增强学习	106
10.2.4 多 Agent 增强学习	107
10.3 仿真实验	107
10.4 算法特点	109

第 11 讲 流形学习	111
11.1 算法原理	111
11.1.1 ISOMAP	112
11.1.2 LLE	113
11.1.3 LE	113
11.1.4 HE	115
11.2 算法改进	115
11.2.1 LPP	116
11.2.2 MFA	117
11.3 算法仿真	119
11.4 算法特点	123
第 12 讲 RBF 学习	126
12.1 基本原理	126
12.1.1 基于 RBF 函数的内插方法	126
12.1.2 RBF 神经网络	129
12.1.3 数据中心的计算方法	130
12.2 算法改进	132
12.2.1 针对完全内插问题的改进方法	132
12.2.2 针对不适宜问题的改进方法	133
12.2.3 广义 RBF 神经网络	134
12.3 仿真实验	134
12.3.1 基于高斯函数的 RBF 学习	134
12.3.2 RBF 学习算法流程	135
12.4 算法特点	136
第 13 讲 稀疏表示	138
13.1 基本原理	138
13.1.1 信号稀疏表示	138
13.1.2 贪婪求解算法	140
13.1.3 凸优化求解算法	141

13.2	算法改进	142
13.2.1	组合 Lasso (Group Lasso)	142
13.2.2	混合 Lasso (Fused Lasso)	143
13.2.3	弹性网络 (Elastic net)	143
13.3	仿真实验	143
13.3.1	OMP 算法	143
13.3.2	APG 算法	144
13.3.3	基于稀疏表示的人脸识别	145
13.4	算法特点	147
13.4.1	算法优点	147
13.4.2	算法缺点	147
第 14 讲	字典学习	149
14.1	基本原理	149
14.2	算法改进	151
14.2.1	最优方向法 (MOD)	151
14.2.2	K-SVD 法	151
14.2.3	在线字典学习法	151
14.3	仿真实验	152
14.3	基于字典学习的视频图像降噪方法	153
14.4	算法特点	154
14.4.1	算法优点	154
14.4.2	算法缺点	155
第 15 讲	BP 学习	156
15.1	基本原理	156
15.1.1	人工神经网络	156
15.1.2	BP 学习原理	157
15.2	算法改进	162
15.2.1	改进学习速率	163
15.2.2	改进训练样本	164
15.2.3	改进损失函数	164

15.2.4	改进连接方式	165
15.3	仿真实验	165
15.4	算法特点	167
第 16 讲	CNN 学习	170
16.1	基本原理	170
16.1.1	神经认知机模型	170
16.1.2	CNN 算法思想	171
16.1.3	CNN 网络结构	171
16.1.4	CNN 网络学习	174
16.2	算法改进	178
16.2.1	设计新的卷积神经网络训练策略	178
16.2.2	使用 GPU 加速卷积运算过程	178
16.2.3	使用并行计算提高网络训练和测试速度	179
16.2.4	采用分布式计算提高网络训练和测试速度	179
16.2.5	硬件化卷积神经网络	179
16.3	仿真实验	179
16.3.1	卷积神经网络训练算法仿真	179
16.3.2	卷积神经网络实际应用实例	181
16.4	算法特点	183
16.4.1	算法优点	183
16.4.2	算法缺点	183
第 17 讲	RBM 学习	185
17.1	基本原理	185
17.1.1	RBM 学习思想	185
17.1.2	RBM 模型基础	186
17.1.3	RBM 模型学习	189
17.2	算法改进	195
17.2.1	方差 RBM	195
17.2.2	均值方差 RBM	196
17.2.3	稀疏 RBM	196