

模式识别

课程简介 &
模式识别系统简介

吴建鑫

南京大学计算机系&人工智能学院，2020

课程介绍

新冠疫情特殊情况下的应对措施

- ✓ 在线课程（学校通知正式复课之前）
 - 教学立方 <https://teaching.applysquare.com/>
 - 邀请码 594046 （2月22日之前有效）
 - 课件：视频见教学立方 & PPT及其他见本课程主页
 - 请使用视频、课件自学
 - 作业、小测试
 - 在课程立方中进行，将通过课程立方提前通知（微信公众号）
 - 交互、答疑
 - 将在课程立方讨论区及QQ群（群号783269697）进行
 - 视情况不定期交互，会在课程立方提前通知
 - 联系： wujx2001@nju.edu.cn

关于选课、补选

- ✓ 目前选课规模约在200人左右
 - 计算机系本科生、研究生；人工智能学院研究生；少量全校公选名额
 - 本人还需要准备开一门新课（全部从头开始准备），无力扩大课程规模，无法再增加选课人数
 - 按学校规定正常退选、补选，**请勿私人联系选课事项**
- ✓ 未能选课，但是感兴趣怎么办？
 - **欢迎在教学立方（及恢复正常教学后）跟班学习**
 - 欢迎在QQ群及教学立方与本人及助教进行交互
 - 助教人数少，无力批改不在花名册同学的作业、试卷
 - **可以学到知识，但是将不能获得学分！**

课程目标

- ✓ 理解、记忆模式识别中的基本概念、步骤和方法
- ✓ 对重要方法，要能实际应用，并能理解其前提条件、应用范围、应用注意事项和方法原理及推导
- ✓ 对讲授的其他方法，要能理解其含义和使用环境
- ✓ 要对模式识别的前沿领域有感性的认识
- ✓ 提高目标
 - 进一步能通过独立阅读和实践掌握较深入的问题和方法，并能应用到学习、研究中遇到的问题中去

模式识别

- ✓ 本科生和研究生通选的选修课，2学分，2学时
- ✓ 前置课程(prerequisite)
 - 程序设计、高等数学（数学分析和线性代数）、概率和数理统计
 - 一般来说，最早2年级下半学期才可以选
- ✓ 没有或者(自愿)1次大作业（视疫情发展定）
 - 有若干小作业
 - 包含纸上解题和一些小的编程(需要Linux软件安装)
- ✓ 成绩考核以考试和作业为主（视疫情发展定）
 - 期末考试50-60%，作业40-50%
 - 自愿做大作业的本科同学有适当加分（视疫情发展定）

教材

✓ 英文

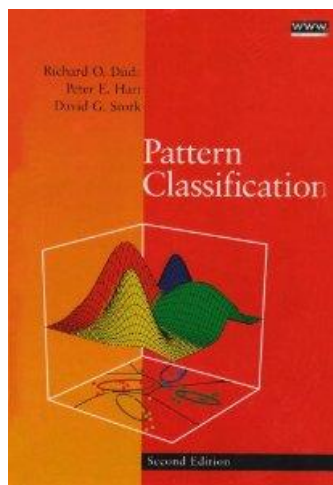
- Essentials of Pattern Recognition - An accessible approach, Jianxin Wu
- 剑桥大学出版社（英国），已定稿
- 目前仍以讲义的形式分章贴在课程主页上

✓ 《模式识别》

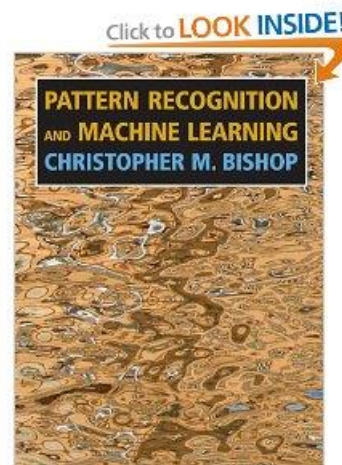
- 吴建鑫著，罗建豪、张皓译，吴建鑫审校
- 机械工业出版社，书号：ISBN 978-7-111-64389-0
- 目前疫情期间情况特殊，暂时还未能购买
 - 请先利用英文版本，待可以购买后会及时通知大家
- 中英文教材欢迎纠错，email联系确认后可加分

参考书

DHS



PRML



参考书中译本: <http://item.jd.com/10057422.html>

Matlab

- ✓ 本课程中用到的编程，很多在Matlab环境中最方便
 - <http://cn.mathworks.com/help/matlab/>
 - Matlab简明教程：请自行搜索Matlab的tutorial
 - 对从事计算机研究几乎是必备的技能
 - 学校已购买Matlab，本校师生均可免费使用（需要校园网）
- ✓ 如果没有Matlab环境
 - 有很多免费软件，实现Matlab的功能足够本课程使用
 - 语法和编程语言与Matlab一致
 - 如GNU Octave：
<http://www.gnu.org/software/octave/>
- ✓ 学会随时调用Matlab的帮助（或在线帮助）

数学背景

- ✓ 会简要介绍/回顾需要的最基本知识
 - 前置课程，（如较好的掌握了）就已经够本课程使用了
 - 要用到的更多的数学知识会在授课中随时补充
- ✓ 对研究或工作有帮助
 - 特别是对有志CS研究的（不限于模式识别），有较大帮助
- ✓ 如对进一步的知识感兴趣
 - 每章会提供关于进一步阅读的资源指南
 - 通用资源：
 - 自编讲义
 - 教科书（DHS），参考书（PRML）
 - 维基（Wiki）
 - Wolfram MathWorld (<http://mathworld.wolfram.com/>)

模式识别简介

什么是模式识别？

- ✓ “Pattern recognition is a field in machine learning.” (??)
 - 取自2013年12月25日，2016年2月9日仍然如此
 - Pattern recognition (**psychology**), identification of faces, objects, words, melodies, etc.
 - [http://en.wikipedia.org/wiki/Pattern_recognition_\(disambiguation\)](http://en.wikipedia.org/wiki/Pattern_recognition_(disambiguation))
- ✓ In machine learning, pattern recognition is the assignment of a label to a given input value.
 - http://en.wikipedia.org/wiki/Pattern_recognition
- ✓ PRML: automatic discovery of regularities in data ... and with the use of these regularities to take actions such as classifying the data into different categories

模式识别的特征

✓ 输入：数据

- data, input value, ...

✓ 输出：“模式” (Pattern)

- faces, objects, words, melodies,
(classification) label, categories,
regularity, ...

✓ 输入转换到输出：

- 由计算机系统完成 (automatic, algorithm, assignment ...)
- 是一个困难但有趣的过程 (discovery)

模式识别举例：Autopilot



✓ 输入？

✓ 输出？

✓ 困难？

视频来源

<https://www.tesla.cn/models>

Autopilot主页

<https://www.tesla.cn/autopilot?redirect=no>

模式识别举例：Kinect



- ✓ 输入？
- ✓ 输出？
- ✓ 困难？

视频来源

http://v.youku.com/v_show/id_XMTgxNDk00Dc2.html

模式识别举例：Siri



- ✓ 输入？ 视频来源
http://v.youku.com/v_show/id_XMzMlMjc1MTgw.html
- ✓ 输出？ Siri的产品主页
<https://www.apple.com.cn/siri/>
- ✓ 困难？

从数据到模式 From data to pattern

传感器

- 产生物理读数
- 物理学、声学、电子学、光学、触觉、味觉、...

数字信息

- 文本、图像、音频、视频、三维点集、...

模式

- 物体检测和识别、音频转化为文本、行为识别、姿态识别、...

应用

- 决策、计划、处理、...

(虚拟例子) 性别识别：特征

✓ 选取什么 **数据** 来判别一个人的性别？

- 长相？喉结？
- 身高？
- 体重？
- 声音？
- 行为？
- ...



✓ 被 **使用** 的数据或从原始输入数据 (raw input data) 中提取 (extract) 的数据称为 **特征** (feature)

- 为什么要选取、提取、或者 **学习** 特征？

(虚拟例子) 性别识别：评估

- ✓ 怎么知道模式识别的结果好坏？
 - 评估或评价(evaluate)
 - 通常是把模式识别系统的输出(称为预测, prediction)与真实值(groundtruth)进行比较
- ✓ 真实值从哪里来？
 - 女性的性染色体 $XX \rightarrow X$ (来自母亲) + X (来自父亲)
 - 男性的性染色体 $XY \rightarrow X$ (来自母亲) + Y (来自父亲)
- ✓ 如何进行比较？
 - 以后会讲。现在先想一想？

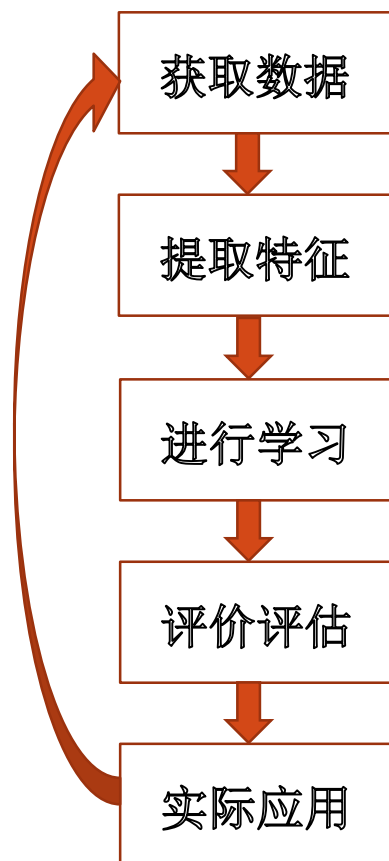
(虚拟例子) 性别识别：转换

- ✓ 如何构建一个模式识别系统将输入转换为输出
 - 常见方法：机器学习(machine learning)
 - 从过去的经验学习Learning from experience
 - 训练集(training set)：已搜集的数据
 - 模型(model)：从训练集得到的规律，表现形式多样
 - 测试(prediction, testing)：将这些规律应用到新的例子以得到针对该例子的结果（模式识别的输出）
 - 测试集(testing set)：为了评估这些规律搜集的数据
- 那么，对训练集和测试集应该有什么要求？

（虚拟例子）性别识别：应用

- ✓ 得到输出后，有什么用？
 - 通常在更大的系统中起作用
 - iOnRoad, Siri, Kinect, ...
- ✓ 那么，模式识别系统要做到多好呢？
 - 由那个“更大的系统”决定
 - 是不是越高越好？

小结：模式识别的步骤



问题：

- ✓ 是这样吗？
 - 需要多次反馈 (feedback)、尝试、修改
- ✓ 什么步骤最重要
 - ??
- ✓ 深度学习在图中什么位置？

模式识别为什么困难？

✓ 多种原因

- 语义鸿沟semantic gap, 如, 一副 2000×1000 的图像是什么?
 - 对人的眼睛和大脑: 教师、桌椅、讲台、黑板、人、...
 - 对计算机: $3 \times 2000 \times 1000 = 6,000,000$ 个数字
- 计算能力computational power
 - 很多算法在台式机上要数年或更长时间
 - 大数据, 存储和计算都成为问题
- 数据的获取data acquiring
 - 很多时候数据难以获取(如, 医学图像medical imaging)
 - 或者虽然容易获取, 但是难以标注(如, 对google图像的精确标注annotation)

www.image-net.org/synset?wnid=n03028079

Numbers in brackets (the number of synsets in the subtree)

- ImageNet 2011 Fall Release (21841)
 - animal, animate being, beast, brute, creature, fauna (3772)
 - plant, flora, plant life (3775)
 - person, individual, someone, somebody (3772)
 - fungus (298)
 - natural object (551)
 - artifact, artefact (7894)
 - paving, pavement, paving material (45)
 - restoration (0)
 - ready-made (0)
 - plaything, toy (17)
 - padding, cushioning (12)
 - opening (18)
 - mystification (0)
 - marker (2)
 - line (26)
 - sphere (0)
 - square (0)
 - building material (41)
 - weight (3)
 - way (68)
 - track (2)
 - thing (2)
 - surface (37)
 - structure, construction (1238)
 - airdock, hangar, repair shed (0)
 - corner, quoin (0)
 - cross (0)
 - deathtrap (0)

Place of worship, house of prayer, house of God, house of worship

Cathedral Cathedral Kirk Abbey Basilica

www.image-net.org/synset?wnid=n03028079#panel

© 2010 Stanford Vision Lab, Stanford University, Princeton University support@image-net.org Copyright Intelligemix

一个例子

- ✓ ILSVRC2010: Large Scale Visual Recognition Challenge 2010
 - 图像分类: 1,000个类别, 1,200,000训练图像
 - 原始输入(图像)约100+GB
 - 有效的图像表示: 每个图像表示为262,144的向量, 需要1258GB空间来存储特征
 - 设计新的机器学习算法, 使得存储开销缩小到约40G
 - 用SVM^{light} 或LIBSVM要很多年
 - 用LIBLINEAR要大概2个月
 - 使用新的深度学习(deep learning)方法, 用GPU辅助计算, 可以得到更好的精度accuracy
 - 尝试一下用你的笔记本电脑**测试**这个网络?

各步骤的（实践）重要性

1. 提取特征

- 设想一下，告诉你一个人的年纪，需要判别其性别？
- 或者，告诉你一个人的染色体组成，需要判别其性别？

2. 数据获取

- 设想一下，根据图像判断性别，所有训练图像都只包含男子、而不包含任何女子？
- 或者，图像中有男有女，但是所有都标记为女？

3. 机器学习

- 有很多理论上重要的问题，对实践效果也非常重要

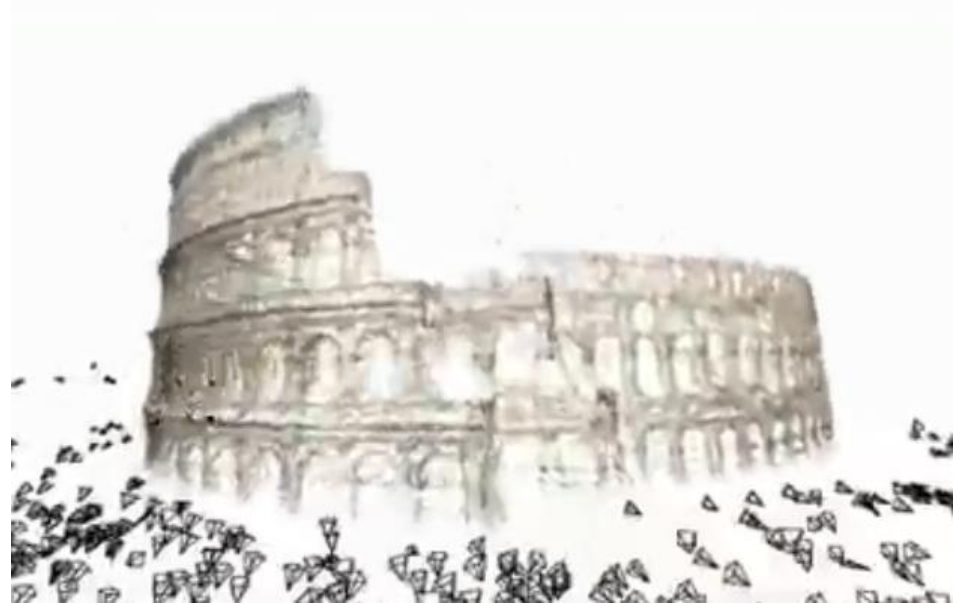
和其他领域的关系

模式识别vs. 机器学习

- ✓ 机器学习在模式识别中有非常重要的作用
- ✓ 但是，模式识别具有更多的“系统”性
 - 数据获取
 - 提取特征
 - ...
- ✓ 主要的区别是：“数据” vs. “特征”
 - 但是，在深度学习中，这个区别不明显了

模式识别vs. 计算机视觉

- ✓ 模式识别与计算机视觉 (computer vision) 的研究和应用有非常多的重合
 - 识别recognition是计算机视觉中最重要的问题之一
 - 模式识别中很大部分输入是图像
- ✓ 模式识别包含很多视觉以外的问题
 - 音频、雷达、文本、...
- ✓ 计算机视觉包括很多识别以外的问题
 - 如，超分辨率super-resolution
 - 三维重建3D reconstruction



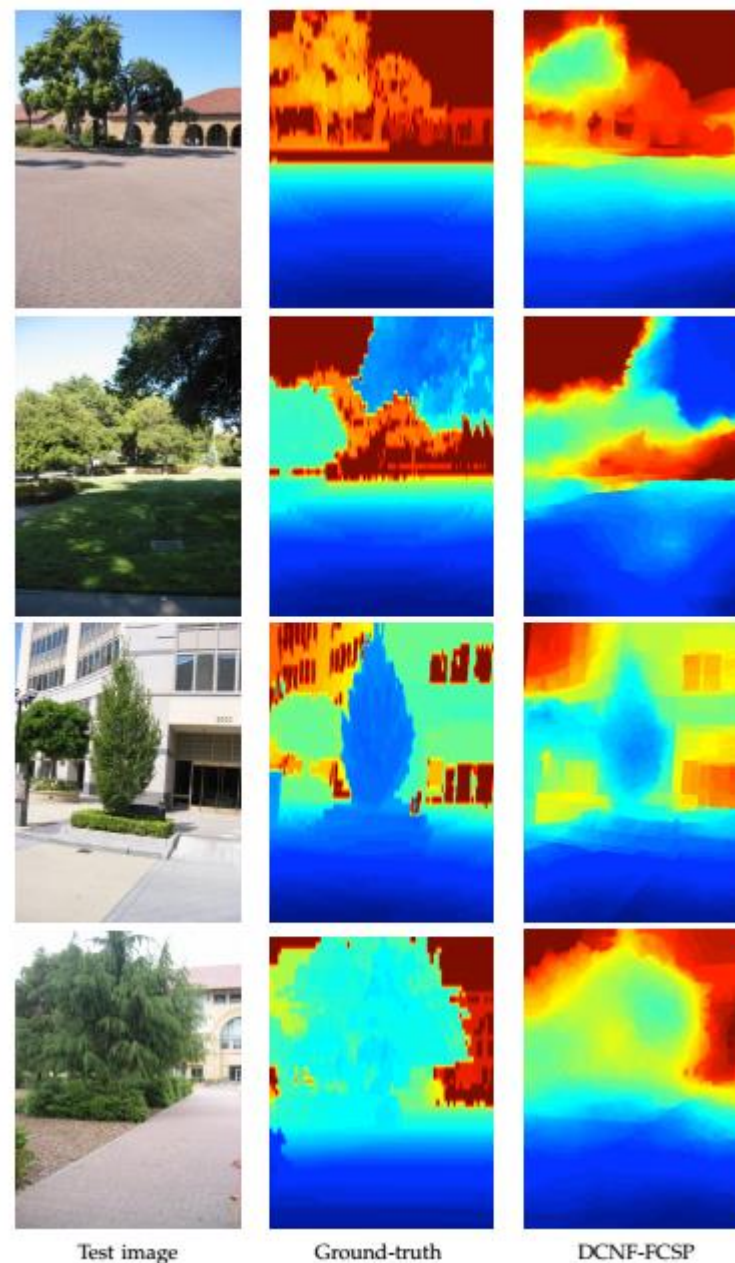
[http://upload.wikimedia.org/wikipedia/commons/e/ed/An example of super resolution with still RAW photo.jpg](http://upload.wikimedia.org/wikipedia/commons/e/ed/An_example_of_super_resolution_with_still_RAW_photo.jpg) (链接已失效)

http://v.youku.com/v_show/id_XMTUwMDg00TQw.html
<http://grail.cs.washington.edu/rome/>

又一个例子

✓ 单张图片深度学习深度估计

- Depth estimation using deep learning from single image
- Learning Depth from Single Monocular Images Using Deep Convolutional Neural Fields
- <http://arxiv.org/abs/1502.07411>
- IEEE T-PAMI
- 阿德莱德视觉组



模式识别与多媒体 (multimedia)

- ✓ 都可能牵涉多种媒体
- ✓ 但多媒体更具有“系统”性
 - 多媒体不特别强调单个模块的性能
 - 更注重整个大系统的成功
 - 比模式识别更强调多种媒体之间的配合
 - 就算每个模块都采用了已有的技术，但是一个科学利用现有技术和多种媒体的系统仍然可以是很大的成功

小结：模式识别vs. 其他学科

- ✓ 模式识别与很多学科有广泛的联系
 - 机器学习、计算机视觉、多媒体、可视化 (visualization)、图形学 (graphics)、...
- ✓ 需要很多学科的支持
 - 物理学、电子学、...
- ✓ 模式识别与相关学科是有区别的
 - 但是，现在各个学科呈现越来越融合的趋势
 - 所以，不要强调区别，让我们更注重共性、合作、以及更重要的一**解决问题**！

进一步的阅读

- ✓ 各相关领域的最新进展可以参考其重要会议和期刊
 - 机器学习: JMLR, ICML, NIPS
 - 计算机视觉: TPAMI, IJCV, CVPR, ICCV, ECCV
 - 多媒体: ACM Multimedia, TMM
- ✓ 有用的资源:
 - 计算机视觉、图像处理OpenCV:
<http://opencv.org/>
 - Matlab: 若干工具箱toolbox
 - 深度学习软件包（百花齐放）！

你适合本课程吗？

- ✓ 拿到学分不容易
 - 需要投入一些时间, 仅课堂听一下不能过关
- ✓ 尤其适合于对研究感兴趣的同学
 - 不一定毕业以后要做研究
 - 你觉得解决一个模式识别问题有意思
 - 想想我们举的几个例子
 - 你觉得模式识别本身挺有意思
 - 或者其他让你觉得有趣的东西
- ✓ 如果只是为了学分, 请退选！