



数字媒体技术

第1章 数字媒体技术概述

刘绍辉 刘贤明

计算机科学与技术学院 哈尔滨工业大学

shliu@hit.edu.cn

2021年春季



自我介绍

姓名：刘绍辉



单位：哈工大. 计算机科学与技术学院. 智能接口与人机交互技术研究中心

研究方向：图像、视频处理，
计算机视觉，模式识别

联系方式：

13503627854, shliu@hit.edu.cn

微信号：shliu13503627854

办公室：综合楼613

课程QQ群号:741407461



群名称：数字媒体技术

群 号：741407461



0.多媒体技术课程简介

- ◆总学时72，讲课：48，实验：24
- ◆上课内容
- ◆考试方法
- ◆课程要求
- ◆参考文献





0.1 上课内容

共计72学时，第1周-第16周，其中48学时上课，24学时实验.上课：1-16，周5，5-6，致知14; 1-8周(双周)，周3，7-8，致知14。实验：1-16周（双周），周2，1-2，实验,格物207

- ◆第1章 绪论(2学时)
- ◆第2章 数字媒体形成的理论基础及度量(6学时)
- ◆第3章 数字媒体的表示(4学时)
- ◆第4章 项目交流与研讨(4学时)
- ◆第5章 媒体的数学变换(6学时)
- ◆第6章 音频和图像的基本处理(4学时)
- ◆第7章 数字媒体的压缩(6学时)



0.1 上课内容(续)

共计72学时，第1周-第16周，其中48学时上课，24学时实验。

- ◆第8章 典型的H.264与HEVC编码平台的编解码(4学时)
- ◆第09章 典型媒体特征提取(6学时)
- ◆第10章 开源机器学习(4学时)
- ◆第11章 数字媒体分析与处理(6学时)
- ◆第12章 基于开源机器学习平台的媒体分析(4学时)
- ◆第13章 数字媒体安全技术(4学时)
- ◆第14章 媒体安全算法实现(4学时)
- ◆第15章 数字媒体存储与检索(4学时)
- ◆第16章 项目讨论与交流(4学时)

0.2 考试方法

◆最终总成绩包括三个部分

- 期末考试成绩40%
- 实验成绩占40%
- 交流与沟通占20%

◆考试方式

- 闭卷考试
- 实验项目开题和结题报告

- ✓ 实验为设计和实现一个多媒体处理系统/多媒体计算系统。整个系统可以分解为：（1）框架设计；（2）数据获取与存储；（3）数据处理与分析；（4）结果展示；（5）实验评估。例如，设计并实现一个多媒体智能教室，采用带云台的多个摄像头，其功能可包括：可进行人脸识别与验证；可进行人数计数；可跟踪教师；可进行录像；可在线提问；可线下提问，摄像头可自动对准提问的同学并进行识别；可远程接入学习；可通过多种媒体进行检索。

◆考试时间地点

- 教务处安排

项目选题：注意利用前序课程《移动互联网技术》中的实验平台

- ◆ 实现一个多媒体智能教室，在教室中，有多个摄像头，可以形成教室的全景图，可以录制教师上课内容的视频和声音，并对学员提问进行响应和回答；
- ◆ 无人监考系统：通过摄像头和麦克风，对考场进行监控，并对考生的姿态和行为进行判断和分析，及时报警提示，并录制监控视频；
- ◆ 实现一个校园出入口车流、人流流量统计系统，并尝试识别车牌号和人脸识别，如果认识，给出语音提示或欢迎提示
- ◆ 设计并实验一个类似抖音、快手等的直播系统，并添加各种基本功能(可基于nginx)

大项目要体现：移动性、工程性、系统性，尤其是软件工程里面的核心问题



0.3 课程要求

数字媒体技术是一门软件工程大类的专业核心课，不仅要求掌握基本的数字媒体相关的获取、存储、处理、分析及应用，而要求，实验性的课程，要求动手做实验

◆ 交流与沟通20%

- 实验项目开题与结题，平时讨论，组内讨论，作业等

◆ 实验成绩40%

- 项目开题和结题

- ✓ 通过课程中间的各个实验内容，最终达到项目的目标，提供设计报告和源代码，并讲解、演示项目成果

- 文献阅读报告

- ✓ 根据项目或与媒体技术相关的最新文献范围(上课当年发表的期刊或者会议文献，主要从相关的顶级会议和期刊中选取)，仔细阅读后将主要思想和文章的贡献写成读书报告上交

◆ 考试成绩40%

- 掌握基本的概念和理论，能对数字媒体技术的各个部分有比较清楚的掌握

◆ 选修同学：考试50%+实验20%+阅读报告10%+交流与沟通20%



0.3课程目标

- ◆能够设计并实现基于文本、语音、图像和视频等多媒体，包括获取、存储、处理、分析和展示的计算系统。掌握在给定的资源情况下，，尤其在移动互联网上，应该如何选择关键设备的技术指标来完成特定的多媒体计算系统。
- ◆掌握用于解决多媒体计算系统设计与实现的关键技术、方法与工具。具体包括：语音、图像、视频获取的基本原理，压缩存储以及表示的基本方法，处理与分析的关键技术。
- ◆具备对多媒体计算相关的复杂工程问题进行分析 and 评价的能力。主要包括问题的分析与评价、算法的分析与评价以及系统性能分析与评价的能力，使学生能够利用数学知识对关键技术进行分析与评价的能力。
- ◆能够对复杂的与媒体相关的实际问题利用《移动互联网技术》中的知识给出综合解决方案，能够撰写与多媒体处理系统相关的论文阅读报告、实验开题与结题报告，并具有对实验内容进行交流与沟通的能力，具备自学、独立思考和一定的创新和跨学科交流的能力。



0.4 参考文献

- ◆ 林福宗, 多媒体技术基础(第3版), 清华大学出版社, 2009.1
- ◆ Yihong Gong, Wei Xu, Machine Learning for Multimedia Content Analysis. Springer, 2007
- ◆ Lakhmi C. Jain. Machine learning for audio, image and video analysis - theory and application. Second version. Springer, 2015.
- ◆ Alan C. Bovik. Handbook of Image and Video Processing. Academic Press., 2000.
- ◆ 白金榜等译, 多媒体原理, 电子工业出版社, 2003.8
- ◆ Wang Yao, Video processing and communication, 清华大学出版社, 2003.2
- ◆ Yun Q. Shi, Huifang Sun. Image and Video Compression for Multimedia Engineering Fundamentals, Algorithms, and Standards(Second version), CRC Press, 2008
- ◆ 历届的相关国际会议论文集, 了解最新的研究动态, 例如ACM Multimedia(顶级国际会议), CVPR, ICCV, ECCV, AAI, IJCAI, ICME, ICASSP, ICIP ..., IEEE Transactions on Multimedia, IEEE Transactions on Image Processing, Pattern Analysis and Machine Intelligence (TPAMI), IJCV...





0.4 参考文献

- ◆ 历届的相关国际会议论文集，了解最新的研究动态，例如ACM Multimedia(顶级国际会议)，CVPR，ICCV，ECCV，ICIP，ICME，ICASSP...，IEEE Transactions on Multimedia，IEEE Transactions on Image Processing，Pattern Analysis and Machine Intelligence (TPAMI)...
- 网上文献查找：学校图书馆的数据库
- 中文：中国知网；万方数据
- 英文：ACM美国计算机学会；Elsevier Science；IEEE/IEE Electronic Library；Springer LINK 电子期刊
- 网络预印本论文：<http://arxiv.org>，网上搜索引擎：google scholar, bing, ResearchGate等，<http://www.cvpapers.com>
- ◆ 今年阅读文献和讲解文献推荐考虑
- AAAI2020：NewYork, USA, Feb. 7-12, 2020
- CVPR2020：Seattle, USA, June 13-19, 2020
- IJCAI2020：Yokohama, Japan, July 11-17, 2020
- NeurIPS2019：Vancouver, Canada, Dec.8-12, 2019
- ACM MM 2019：Nice, France, 21-25 Oct. 2019
- CVPR2019：Long Beach, CA, USA, June 16th –June 20th
- ICCV2019：Seoul, Korea, Oct.27-Nov.2, 2019
- 也可选下面2018的会议论文
- ACM MM 2018 :October 22-26, 2018, Seoul Korea
- CVPR2018的论文：June 18-22, 2018, Salt Lake City USA
- ICCV2018的论文：Jan 30-31, 2018, Istanbul, Turkey

更多相关会议参考：

<http://www.mamicode.com/info-detail-2796300.html>

提示：

目前预印本网站 <https://arxiv.org>上有很多投递但未审稿的最新论文，大家可以从其中阅读并报告，例如目前还没有开的会议，但是投递出去的论文在这个网站和作者主页上能找到不少！

0.5 国内动态：中央政府高度重视

- ◆ 2017年07月20日，国务院发布《新一代人工智能发展规划》（国发〔2017〕35号）
- ◆ 《规划》经过中央政治局常委会、国务院常务会议审议



中华人民共和国中央人民政府
www.gov.cn

国务院 总理 新闻 政策 互动 服务 数据 国情

首页 > 信息公开 > 国务院文件 > 科技、教育 > 科技

索引号: 000014349/2017-00142
发文机关: 国务院
标 题: 国务院关于印发新一代人工智能发展规划的通知
发文字号: 国发〔2017〕35号
主 题 词:

主题分类: 科技、教育\科技
成文日期: 2017年07月08日
发布日期: 2017年07月20日

国务院关于印发
新一代人工智能发展规划的通知
国发〔2017〕35号

各省、自治区、直辖市人民政府，国务院各部委、各直属机构：
现将《新一代人工智能发展规划》印发给你们，请认真贯彻执行。

国务院
2017年7月8日
(此件公开发布)

相关报道
* 国务院印发《新一代人工智能发展规划》

图解
* 国务院印发《新一代人工智能发展规划》

解读

0.5 新一代人工智能发展规划

国务院关于印发
新一代人工智能发展规划的通知
国发〔2017〕35号

各省、自治区、直辖市人民政府，国务院各部委、各直属机构：
现将《新一代人工智能发展规划》印发给你们，请认真贯彻执行。

国务院
2017年7月8日

（此件公开发布）

◆ 战略态势

新一代人工智能发展规划

- 我国发展人工智能具有良好基础。国家部署了智能制造等国家重点研发计划重点专项，印发实施了“互联网+”人工智能三年行动实施方案，从科技研发、应用推广和产业发展等方面提出了一系列措施。经过多年的持续积累，我国在人工智能领域取得重要进展，国际科技论文发表量和发明专利授权量已居世界第二，部分领域核心关键技术实现重要突破。语音识别、视觉识别技术世界领先，自适应自主学习、直觉感知、综合推理、混合智能和群体智能等初步具备跨越发展的能力，中文信息处理、智能监控、生物特征识别、工业机器人、服务机器人、无人驾驶逐步进入实际应用，人工智能创新创业日益活跃，一批龙头骨干企业加速成长，在国际上获得广泛关注和认可。加速积累的技术能力与海量的数据资源、巨大的应用需求、开放的市场环境有机结合，形成了我国人工智能发展的独特优势。

0.5 新一代人工智能发展规划



◆战略目标

- **2020年**，人工智能总体技术和应用与世界先进水平同步，人工智能产业成为新的重要经济增长点，人工智能技术应用成为改善民生的新途径，有力支撑进入创新型国家行列和实现全面建成小康社会的奋斗目标。
- **2025年**人工智能基础理论实现重大突破，部分技术与应用达到世界领先水平，人工智能成为带动我国产业升级和经济转型的主要动力，智能社会建设取得积极进展。
- **2030年**人工智能理论、技术与应用总体达到世界领先水平，成为世界主要人工智能创新中心，智能经济、智能社会取得明显成效，为跻身创新型国家前列和经济强国奠定重要基础。





0.5 新一代人工智能发展规划

◆重点任务

➤构建开放协同的人工智能科技创新体系

- ✓建立新一代人工智能基础理论体系：突破应用基础理论瓶颈。瞄准应用目标明确、有望引领人工智能技术升级的基础理论方向，加强大数据智能、跨媒体感知计算、人机混合智能、群体智能、自主协同与决策等基础理论研究。

1. 大数据智能理论。研究数据驱动与知识引导相结合的人工智能新方法、以自然语言理解和图像图形为核心的认知计算理论和方法、综合深度推理与创意人工智能理论与方法、非完全信息下智能决策基础理论与框架、数据驱动的通用人工智能数学模型与理论等。

2. 跨媒体感知计算理论。研究超越人类视觉能力的感知获取、面向真实世界的主动视觉感知及计算、自然声学场景的听知觉感知及计算、自然交互环境的言语感知及计算、面向异步序列的类人感知及计算、面向媒体智能感知的自主学习、城市全维度智能感知推理引擎。



0.5 新一代人工智能发展规划



◆重点任务

- 构建开放协同的人工智能科技创新体系
- 建立新一代人工智能关键共性技术体系
 - ✓ 跨媒体分析推理技术。重点突破跨媒体统一表征、关联理解与知识挖掘、知识图谱构建与学习、知识演化与推理、智能描述与生成等技术，实现跨媒体知识表征、分析、挖掘、推理、演化和利用，构建分析推理引擎。

2. 跨媒体分析推理技术。研究跨媒体统一表征、关联理解与知识挖掘、知识图谱构建与学习、知识演化与推理、智能描述与生成等技术，开发跨媒体分析推理引擎与验证系统。

8. 自然语言处理技术。研究短文本的计算与分析技术，跨语言文本挖掘技术和面向机器认知智能的语义理解技术，多媒体信息理解的人机对话系统。



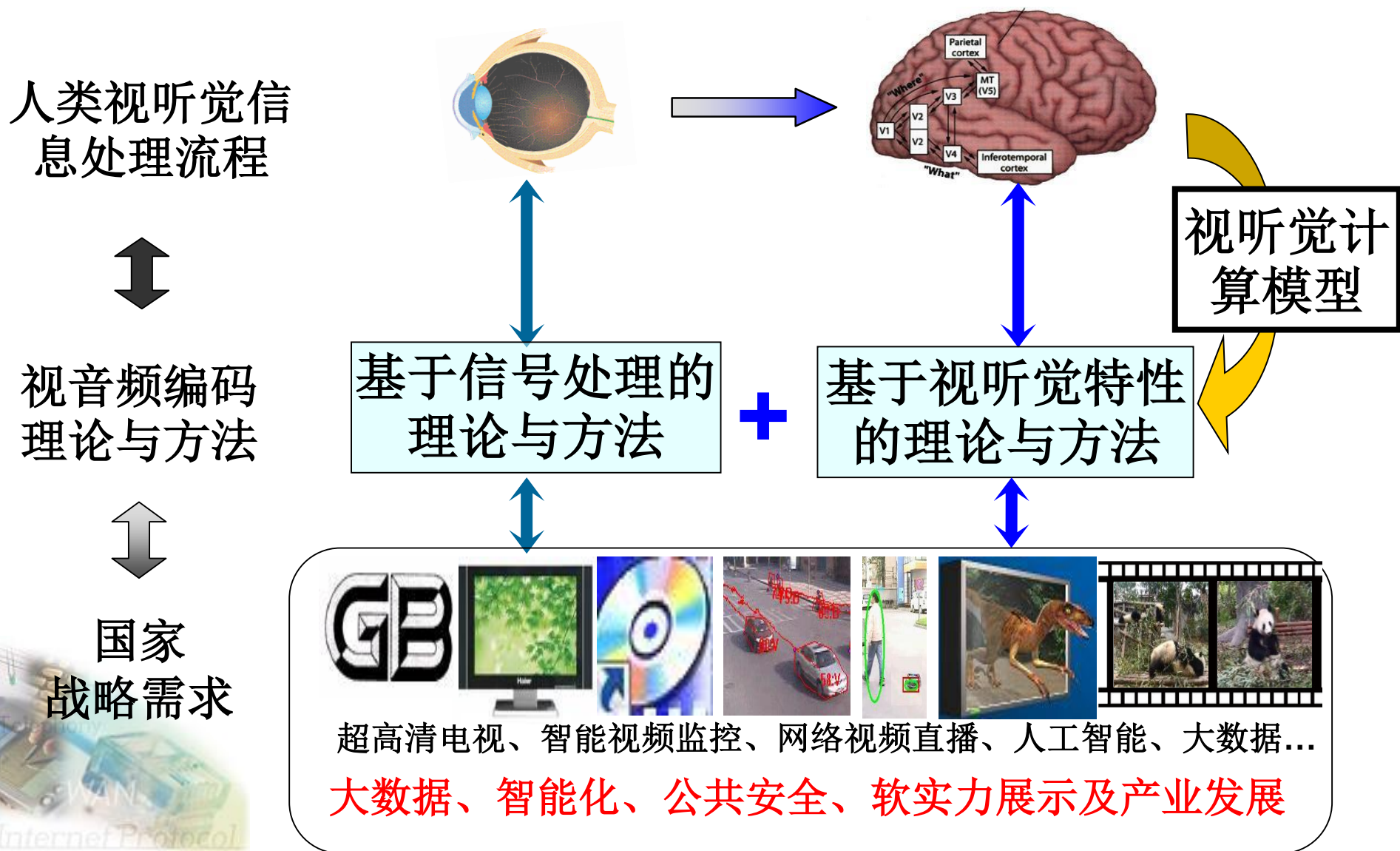


0.5 新一代人工智能发展规划

◆ 培育高端的智能经济、加快推进产业智能化升级、建设安全便捷的智能社会、加强AI领域军民融合、构建泛在安全高效的智能化基础设施体系

		具体领域		
培育高端的智能经济	大力发展人工智能新兴产业	智能软硬件	智能机器人	智能运载工具
		虚拟现实与增强现实	智能终端	物联网基础器件
	加快推进产业智能化升级	智能制造	智能农业	智能物流
		智能金融	智能商务	智能家居
建设安全便捷的智能社会	发展便捷高效的智能服务	智能教育	智能医疗	智能健康和养老
	推进社会治理智能化	智能政务	智慧法庭	智慧城市
		智能交通	智能环保	
	提升公共安全保障能力	公共安全智能化监测预警与控制体系	智能安防	智能化食品安全预警、自然灾害预警与综合应对平台

0.6 补充材料-视听觉关键科学问题





0.6 补充材料-国家政策

◆ 当前AI的新特征

- 第一，形成了大数据上的深度学习与自我锻炼的综合进化技术，这一点在AlphaGo体现的很明显。
- 第二，基于网络下的群体智能时代即将到来，国外《Science》曾将群智计算按难易程度分为实现任务分配的众包模式、较复杂支持 workflows 模式，以及最复杂的协同求解问题的生态系统模式。
- 第三，人机一体化的技术导向混合智能。
- 第四，**跨媒体推理**从语言、视觉、图形与听觉之间的语义贯通间可以形成更好的构建升级。
- 第五，无人系统崛起迅猛。

◆ 国务院的五大核心集成

- 大数据智能、**跨媒体智能**、群体智能、混合增强智能、自主无人系统
- 从基础理论、支撑体系、关键技术、创新应用四个层面构筑知识群、技术群和产品群的生态环境，抢占人工智能技术制高点，妥善应对可能带来的新问题和新的挑战，促进大众创业万众创新，使人工智能成为智能经济社会发展的强大引擎



0.6 产业界动态:加快产业布局

◆ 产业布局加快



(阿里云公司城市大脑)



(百度公司自动驾驶平台)



(腾讯公司医疗影像平台)



(科大讯飞公司智能语音平台)



- 2017年10月11日, 阿里巴巴集团正式宣布成立“**达摩院**”, 进行基础科学和颠覆式技术创新研究。
- 预计未来三年, 投入**超1000亿元**。
- 研究方向涉及**机器学习、基础算法、视觉计算、自然语言处理、人机自然交互**等人工智能相关领域。
- 此前, 百度、腾讯、科大讯飞先后成立了人工智能研究院或实验室



0.6 商汤科技AI



◆2018.9.20，商汤成为继阿里云公司、百度公司、腾讯公司、科大讯飞公司之后的第五大国家人工智能开放创新平台。



智慧安防



智能终端



互联网娱乐



智慧金融



智慧商业



遥感



移动运营商