

# 视听觉信号处理课程报告

卢兑琬

(哈尔滨工业大学 计算机学院, 哈尔滨 150000, e-mail: nty0725@naver.com)

## 1 学习收获

秋去冬来, 经过 72 个学时对《视听觉信号处理》这一门课程的学习, 在老师的教导之下, 我感到自己的视听觉专业方法得到了加强, 专业技能也得到了提高。

### 1.1 主要学习内容

首先在信号与系统的分析、处理这一章节的内容中, 我掌握了信号与系统的定义以及关系、并掌握了信号种类、采集方式, 在此基础上理解了信号表达与编码方法及其应用。在信号处理方法这一部分知识中则是掌握了傅里叶变换的概念和特性, 以及掌握傅卷积运算的概念和特性。在信号分析中则是能够理解和运用正交概念, 并正确理解信号分解, 进行信号的稀疏表示。在接下来的语音学识别分析内容中, 我理解了语音和其表示, 并从生理学、心理学、声学、语音学、韵律学和语言学等多个学科理解语音产生、传输、感知、认知等各方面的特性, 并在老师的教导下能够理解并综合运用各学科知识来建立各个仿真数学模型, 并理解了语音处理研究的发展轨迹。在语音的时域频域分析部分, 我理解了特征分析在语音处理研究中所扮演的重要角色、理解了各类模型和数据工具的运用技巧、掌握了各种时频分析方法, 并且能够在典型语音信号处理问题中选择和实现适当的分析算法。在语音编码技术的学习部分, 首先了解语音编码的概念和发展轨迹, 并在此基础上理解了语音编码技术的基本框架、理解了主要的语音编码算法、了解了最新研究动态, 并最终能够设计和实现典型的语音编码算法。在语音部分的最后内容—语音技术识别概述一节中, 了解了语音识别任务和其对应的不同解决方案, 并了解了说话人识别任务和其多种解决方案。在数字成像、光度学与色度学部分的学习内容中, 我掌握了数字成像(包括电磁波成像、声波成像)特点、了解了图像处理历史发展状况, 并掌握光度学、色度学及视觉特性, 在前述知识的基础上理解了彩色空间、形状、纹理、尺度关系。在空域图像增强方法部分的学习, 在老师对各类空域图像平滑算子以及各类空域图像锐化算子的讲解下, 我理解了图像处理中各类算子的作用、理解了各类模型和数据工具的运用技巧、掌握了各种图像处理算子方法, 并具有一定的科研动手能力, 能够在指导下设计实现针对问题的图像特征提取系统。在接着的频域图像增强方法学习部分, 我理解了频域图像处理中各类算子的作用、理解了各类模型和数据工具的运用技巧、掌握了各种图像处理算子方法, 并在专业知识的基础之上具有一定的科研动手能力, 能够设计实现针对问题的图像特征提取系统。在最后的图像复原小节则是掌握了图像退化模型, 并掌握图像恢复原理及图像复原方法。

在实验方面, 分别了解掌握了语音特征提取实验、语音编码实验、命令词识别实验、

图像读写、彩色空间转换及图像运算变换实验、成图像增强算子实验的基本内涵及操作。在专业知识学习与实验操作方面，基本掌握了视听觉信号处理的基本内容。

## 1.2 本人学习体会

图像、语音和文本（语言）作为与人类视听觉感知密切相关的信息，在社会、经济和国家安全等领域中扮演着重要角色，并在今后一段时间内仍将迅猛增长。这类信息可被人类直接感知和理解，也可用计算机进行处理。虽然如今计算机的处理能力在日益迅速地增长之中，但尚未能够满足当今智能社会的发展需求。如何借鉴人类的认知机理和相关数学的最新研究成果，建立新的计算模型和方法，从而大幅度提高计算机对这类信息的理解能力与处理效率，不仅可有力推动信息科学的快速发展，也将为国民经济和社会发展做出重大贡献。

在学习《视听觉信号处理》这门课程的过程之中，从老师的专业知识讲解以及实验过程操作的学习之中，我感受到了这门课程在围绕着国家的重大需求，充分发挥着信息科学、生命科学和数理科学的交叉优势，从人类的视听觉认知机理出发，研究并构建新的计算模型与计算方法，提高计算机对非结构化视听觉感知信息的理解能力和海量异构信息的处理效率，克服图像、语音和文本（语言）信息处理所面临的瓶颈困难，今后对视听觉信息处理此方向的学习继续深入之后，将会为确保国家安全与公共安全、推动信息服务及相关产业发展以及提高国民生活和健康水平做出重要贡献。具体表现为：在视听觉信息处理的基础理论研究方面取得重要进展；在视听觉信息协同计算、自然语言（汉语）理解以及与视听觉认知相关的脑—机接口等三项关键技术方面取得重大突破；集成上述相关研究成果，研制具有自然环境感知与智能行为决策能力的无人驾驶车辆验证平台，主要性能指标达到世界先进水平，从而提升我国在视听觉信息处理领域的整体研究实力，培养具有国际影响力的优秀人才与团队，为国家安全和社会发展提供相关研究环境与技术支撑。

## 2 视听觉信号处理发展现状

目前对视听觉信号处理的研究热点关注着“感知特征提取、表达与整合”、“感知数据的机器学习与理解”和“多模态信息协同计算”等核心科学问题。

### 2.1 视听觉信息协同计算的关键技术

研究机器视听觉信息的协同计算模型及系统实现技术，基于视听觉信息融合的模式识别技术与验证系统，跨模态视频信息检索与网络敏感信息过滤技术及应用。实现基于多模态协同计算模型的网络视频信息搜索的查准率比同期国外最好水平高 5%—10%，并在网络信息安全与服务等领域得到验证。

### 2.2 与视听觉认知相关的脑—机接口关键技术

研究与视听觉认知相关的脑信号提取、脑区定位与脑功能网络分析技术，脑—机交互中的信号传输、处理、控制技术及系统实现，与视听觉认知相关的脑—机接口典型应用。实现所提无创脑—机接口信息提取与分析等技术处于同期国际领先水平，并能够在改善残疾人生活质量和功能康复等方面得到验证或应用。

### 2.3 无人驾驶车辆集成验证平台

集成上述基础理论与关键技术的相关研究成果，将传统视觉计算模型与新的视觉认知模

型相结合，实现环境感知与建模方法新突破；实现多传感器跨模态跨尺度信息融合，生成高质量三维场景认知地图，构建高性能智能车辆无人驾驶验证平台；提供新的基于人一车一路状态综合分析的智能辅助安全驾驶关键技术；在国防、智能辅助安全驾驶等相关领域得到验证或应用并产生重要影响。

### 3 总结

在长课时的学习过程中，一方面感谢授课老师的辛勤付出，使得我自己的专业水平得到了提高；另一方面，作为留学生不能返校上课的特殊情况，感谢助教老师以及同学对我学业方面的帮助，希望疫情早日结束，尽快重归熟悉的哈工大校园。