

夏立勋

联系方式

江苏省昆山市玉山镇集街

江南新村 09-301

邮政编码: 215300

电话: 13013815427

邮件: lixun.xia@outlook.com

工作经验

- 2014.6 - 至今

主任工程师, 技术负责人, 数字信号处理, 哈曼汽车电子系统(苏州)有限公司, 苏州, 中国

- 为嵌入式音频产品设计和实现基于深度神经网络的轻量级关键词唤醒器(KWS)
- 设计和实现基于深度神经网络的单麦克风降噪模型
- 为嵌入式音频产品设计和实现基于深度神经网络的语音活动检测(VAD)模型
- 设计和开发通用智能音箱产品全自动化测试工具
- 为汽车音频和嵌入式音频产品开发传统信号处理模块
- 音频和语音测量 (全消室, 半消室, ETSI混响室, ACQUA系统等)
- 开发语言和框架: Rust, Python, C++, Julia, Matlab, PyTorch

- 2010.9 - 2014.3

研发工程师, Acosense AB, 歌德堡, 瑞典

- 为基于PLS回归模型的有源声谱流体化学性质实时测量设备开发声学特征模型
- 开发计算声学特征数据的信号处理算法
- 在FPGA平台实现信号处理算法
- 开发仪器信号采集相关的部分硬件平台
- 开发仪器的工业现场总线与SCADA中心通信
- 工业电子产品的CE认证的实施与分析

教育

2008.9 - 2010.9 硕士 集成电路系统设计, Chalmers 大学, 瑞典

2005.0 - 2008.6 硕士 控制科学与工程, 中南大学, 中国

2001.9 - 2005.6 学士 控制科学与工程, 中南大学, 中国

项目经验

- **适用于嵌入式音频产品的轻量级语音活动检测(VAD)系统 (2020 - 至今)**

- 为MIPS和内存受限的电池供电音频产品开发
- 开发了深度模型的整个训练框架
- 目前正在开发当中

传统的VAD工具的一种实现是：首先计算输入信号的各种特征，例如每一帧的MFCC, 周期性和谐波性, 频带的能量等, 然后使用AdaBoost算法进行训练, 训练的结果为最有效特征的重要性排序, 以及每个特征的阈值等信息。这种方法等价于一个单一隐含层的全连接神经网络。受限制与浅层神经网络的性能, 使用深度神经网络能够极大提升分类性能。我们的目标是，在定义了最大的MIPS和内存限制后, 设计和实现一个深度模型的任务训练框架。此框架训练的模型能够有超过工业界常规VAD工具的性能, 如Web RTC的VAD实现(基于Gaussian Mixture Model)。

- **适用于嵌入式音频产品的轻量级关键词唤醒器(KWS) (2019 - 2020)**

- 为便携式音频产品开发通用的可替换关键词的语音唤醒器, 以及唤醒器的任务训练框架
- 为深度模型收集和处理数据
- 特征函数的开发, 模型训练, 模型选择以及验证
- 为BU开发模型推断的演示代码使其能够应用于产品中

基于深度模型的语音唤醒器越来越得到重视，因为随着音频产品的小型化，例如TWS蓝牙耳机，接听电话等操作需要用户抬起手臂到耳朵高度，这在某些情况下是很麻烦的。语音唤醒能够改善用户体验和可用性。然而电池供电系统需要对KWS进行MIPS和内存的限制。此项目目标是优化单位MIPS乘以单位内存的唤醒率指标。语音的领域知识，信号处理技能，以及先进的开发语言的使用，加上收集的海量数据，使得能够在普通PC机器的平台上在有限时间里训练出能够使用的深度模型。

- **三星 Galaxy Home 智能音箱的全自动性能测试 (2018 - 2019)**

- 全自动的测量产品的KWS和ASR在各种复杂条件(噪声，播放回声以及摆放角度等)下的性能指标
- 测试产品在各种认证下的性能指标 (三星认证，微软Cortana和Skype认证, 亚马逊AVS认证以及 Google ART 认证)
- 为各种认证开具最终报告

三星AI部门的合作项目。为其智能音箱产品完成全自动认证程序。Galaxy Home不同于以往的各种产品的一个特点是其多通道AEC系统和基于6个中高音喇叭的具有波束成形的回放能力。这给测试带来了很高的难度。得益于以往开发的通用智能音箱测试软件，高效率的客观测试得以能够在有限的时间内高质量的完成。

- **通用智能音箱产品测量软件 (2018-2019)**

- 自动测试智能音箱产品的语音识别(ASR)以及关键词唤醒(KWS)在噪声，回声等请况下的客观性能
- 在全消室以及混响室涵盖所有测试情况 - 安静/噪声/回声/噪声+回声以及摆放角度等等因素
- 已经用于评估市场部分智能音箱产品 - Alexa Echo/EchoDot, Harman Kardon Invoke, Apple HomePod, Google Home 以及 Samsung Galaxy Home
- 生成的可执行文件体积小，依赖少，内存使用安全可靠，能够长时间运行各种客观测试
- 可扩展能力优秀，甚至能够加入各种未知产品经行黑盒测试
- 无人监督运行，鲁棒可靠

智能音箱的设计开发需要各种信号处理算法的快速迭代，而高效准确的客观测试算法效果对项目成功的重要性不言而喻。基于Rust语言，开发的自动测试软件能够解决此类痛点。其无运行态依赖以及内存安全的特点使得长时间无监督运行成为可能。DUT的状态控制，测试转台的控制，摄像头的控制，以及声压级自动校准等特性使得测量效率极大提升。

- **单麦克风降噪 (2017 - 2018)**

- 收集并处理了哈曼已知规模最大质量最好的噪声数据库

- 基于深度神经网络，业界先进的性能指标(SDR > 13.5dB)
- 模型/数据的自更新能力使得数据库的规模化更加容易

与传统语音增强的降噪算法比较，基于深度网络的算法能够克服更多的噪声情况，比如著名的键盘敲击声。使用单一麦克风意味着所有的信息必须来源于对语音和噪声的特征的理解。这些理解将被训练成模型进行推理。此项目为哈曼嵌入式音频部门开发，同样对模型的MIPS和内存尺寸进行了限制。使用了Mel频谱作为模型的输入；模型的输出为带噪频谱的语音Mask估计。使用Mask分离出语音估计后使用Sqrt Hann窗重新重构到时域以形成干净的语音信号。发现此方法能够对非平稳和暂态噪声有较好性能，对语音的失真影响较小。

- **哈曼卡顿 Invoke 智能音箱语音增强算法调试及测试 (2016 - 2017)**

- 微软远场语音交互AI产品
- 调试麦克风前端处理 (AEC/BF/BS/ABF/NR/Leveller)以完成Cortana以及Skype认证
- 使用ACQUA系统测试
- 哈曼全球杰出项目奖励

微软的语音产品奠定了行业的事实标准。Cortana以及Skype认证是人-机器以及人-人交互的著名标准。我们仔细研究了所有相关标准的内涵和实现，以帮助调试麦克风前端信号处理算法，并完成认证测试。使用了ACQUA系统帮助完成Skype认证。对ACQUA系统的开发和研究促使本人开发了通用的全自动智能音箱测试软件，以弥补ACQUA系统在KWS和ASR的客观测试这方面的不足。

- **代码移植 (Matlab -> C++)：语音去混响 (2015 - 2016)**

- 快速RT60在线估计
- 通过短时间窗口和重叠时间进行直达/混响帧的判断
- 频域到时域的合成过程中实现帧增益控制达到去混响
- 用于实时演示的VST实现

- **代码移植 (Matlab -> C++)：语音自动增益控制 (2015 - 2016)**

- 基于Adaboost的VAD工具
- 通过分析窗/合成窗通过OLA框架在每一帧上进行增益控制，增益控制算法避免了引入过多语音失真
- 用于实时演示的VST实现

- **代码移植 (Matlab -> C++)：基于Adaboost的语音活动检测(VAD) (2015 - 2016)**

- 基于帧的特征提取MFCC，Band Energy，Harmonicity 等，包含一阶和二阶统计
- 能够训练基于Adaboost的模型
- 用于实时演示的VST实现

- **代码移植 (Matlab -> C++)：引擎声降噪 EOC/ANC (2014-2015)**

- 实现了驾驶舱内部的引擎噪声消除
- 多喇叭多麦克风配置
- 控制状态机能够容忍干扰事件比如车窗打开等
- 相关功放产品已经部署到很多北美车型

引擎噪声是发动机转速的基波以及谐波的复杂组合。我们使用了自适应滤波器追踪噪声源到各个乘客之间的传输路径。估计出的传输函数用来卷积喇叭的反相位信号使得在乘客位置声压级最小。我们研究了

Matlab代码，为DSP平台做了C/C++实现。实现了各个支持工具集，比如能够高效准确的测量冲击响应的软件。

- **新型功放电源轨迹跟踪方法** (2014 - 2015)
 - 创新的H级电源轨迹跟踪电路
 - 节省DAC成本
 - 提高了车载功放部门的利润
- **ACOspector(TM) – 非侵入式液体性质实时测量仪** (2010 - 2014)

领导开发了基于主动声谱分析的工业用液体性质实时测量仪器，产品已经部署在瑞典几大造纸和化工企业。具体信息可以参考www.acosense.com。主要贡献包括：

- 开发并生产了第一代和第二代客户能够使用的低成本高可靠的产品
- 产品帮助公司在公共板块上市

- **矩阵式电能变频器** (2007 - 2008)

这是硕士论文项目，国家自然科学基金项目。矩阵式电能变频器能够以极高的能量密度驱动感应/永磁电动机。相比较与传统的背靠背变频器，其代价为需要同时控制9个或者更多的双向电力电子高速开关（IGBT）。我们使用改进的四臂结构能够帮助简化控制任务，并在理论上使其等价与基于载波调制的控制方式。贡献为：开发了第一台原型机。

1. Su Mei, Xia Lixun, Sun Yao, et al "*Carrier modulation of four-leg matrix converter based on FPGA*", Electrical Machines and Systems, 2008. ICEMS 2008. IEEE International Conference on. pp.1247-1250.
2. Yao Sun, Mei Su, Lixun Xia, et al "*Randomized carrier modulation for four-leg matrix converter based on optimal Markov chain*", Industrial Technology, 2008. ICIT 2008. IEEE International Conference on. pp.1-6.
3. Hengsi Qin, Mei Su, Lixun Xia, et al "*A novel controller design method for power converters*", IEEE 11th Workshop on Control and Modeling for Power Electronics, 2008. COMPEL 2008. IEEE International Conference on

语言技能

1. 汉语 - 母语
2. 英语 - 流利

驾照 - C1