Xia Lixun (夏立勋)

> *江南新村 09-301,*
> *集街,*
> *昆山玉山镇, 215300*
> *苏州市, 江苏省, 中国*

> *Mobile Phone: 13013815427*
> *Email: lixun.xia@outlook.com*

Experience

- 2014.6 - present
  Principle Engineer, Dept. Lead, Digital Signal Processing, Core of Competence, Harman Suzhou
  - Machine Learning: Voice Activity Detection(VAD), Single-Mic Noise Reduction, Key-Word Spotting (KWS)
  - Acoustic signal processing
  - Acoustic measurements
  - Rust/Python/C++/Julia/Matlab/PyTorch/Tensorflow
  - FPGA/DSP development
- 2010.9 - 2014.3
  Engineer, R&D, Acosense AB, Göteborg, Sweden
  - Machine learning feature engineering
  - Embedded system design and development
  - VHDL/micro-controller programing
  - Circuit design and PCB layout
  - Knowledge of industrial field-bus
  - Knowledge of CE certification of EE products

Education

2008.9 - 2010.9 *M.Sc.* Integrated Electronic System Design, Chalmers University, Sweden

2005.0 - 2008.6 *M.Sc.* Power Electronics, Central South University, China

2001.9 - 2005.6 *B.Sc.* Control Science and Engineering, Central South University, China

Projects

- ***Light-Weight Key Word Spotter (KWS) for Embedded Audio Devices (2019 -)***
  - Data collection and data processing
  - Training of deep models and benchmarking
  - Model inference and VST

For MIPS and memory constraint embedded audio devices like TWS headhones, key-word spotting must be realized with comparable performance from competitors. We successfully implement a framework for arbitrary KWS to be integrated into embedded audio products based on CNN models.

- ***Tuning/Benchmark Samsung Galaxy Home AI Speaker (2018-2019)***

  - Algorithm tuning for KWS and ASR tasks
  - Benchmark performances based on Samsung internal test standard and Amazon AVS
  - Responsible for final performance and certifacation

  This is my latest project for Samsung AI group. We are responsible for the microphone preprocessing algorithms for Galaxy Home AI speaker. Like Harman/Kardon Invoke, it consists of AEC/BF/BS/NR stages to enable far-field speech recognition, but with the focus on an upgraded multi-channel AEC (M-AEC) which allows for echo cancellation in multichannel playback scenario, given Galaxy Home's 6 tweeters and one woofer. We succeded in tuning the system to meet the global requirement as well as the AVS specification (v3.5.6). As usual my tooling capability also allows for such task completed with minimal people incolved --- only 2 in Suzhou Labotory for this project!

- ***Universal AI Sound Product Test Software (2017-2018)***

  - Automatic test of Automatic Speech Recgonition (ASR) and Keyword Spotting (KWS) performance of all AI products
  - All test cases covering Quiet/Noise/Echo/Echo+Noise in anechoic/ETSI environment
  - Capable of benchmark all AI speech entrance - Alexa Echo/Microsoft Invoke/Apple HomePod/Google Home
  - Written in Julia with excellent runtime performance and scalability
  - Robust enough for operations without human supervision

  This is my latest effort as part of the laboratory instrument automation. Instead of using traditional C++/C# as development language, Julia is chosen for its fast prototyping and high performance. Its parallel/concurrent language models prove itself to be the perfect tool in running asynchronous tasks in such tests. Speech, noise and echo music sound pressure levels (SPL), measured precisely in dBA, can easily be set according to a specification file, allowing the device-under-test to be exposed in precise signal-to-noise ratio during the ASR/KWS tests. Scoring of ASR/KWS test is performed asynchronously to minimize the time cost. This project is product ready and has been constantly used in daily work.

- ***Single-Mic Noise Reduction Based On Deep Neural Network (2017 - 2018)***

  - build of high-quality speech and noise data
  - deep nerual network
  - state-of-the-art Signal Distortion Ratio (SDR > 13.5dB)
  - model bootstrapping
  - machine aided auto-labelling

  This is an on-going project in exploring data-driven single microphone noise reduction. High quality noise data are collected with care for deep nerual net training. Mel band spectrogram as input feature is used to infer the binary/fraction mask from the noisy speech. The mask combined with the spectrogram are eventually synthesized back to time domain to form the clean speech. The signal-to-distortion ratio is comparable to state of the art values. The deep nerual net trained with a small dataset is brilliantly used as VAD tool to expand data labelling.

- ***Tuning/Benchmark of Harman Kardon Invoke (2016 - 2017)***
    - Microsoft far-field AI speakerphone
    - tuning paramters of mic frontend processing (AEC/BS/ABF/NR/Leveller) for ASR, KWS and Skype
    - Hands on ACQUA test for communication
    - Harman global reward for successful project

  Get insight into AI sound products with parameter tuning and performance test. Experience of using ACQUA for acoustic test.

- ***Voice Activity Detection (VAD) based on Decision Tree (2016 - 2017)***
    - Frame features of speech including first/second moment statistics
    - Decision tree models are built upon tabular feature data
    - Implemented in VST for real time demonstration

  The intention of this project is to realize a light weight VAD tool based on carefully chosen audio features found in speech: MFCC/Periodicity/band energy and many more. Decision tree models are considered to allow real time implementation. For clean speech our model has performance comparable to relavent VoiceBox functions. Besides, our model is trained to be working with moderate noises especially transient noise. Unlike deep nerual net this model doesn't require as many data in training if support vectors are sufficient, so it fits many use cases like voice control within car/vehicle and/or ultra low power devices.

- ***Voice De-Reverberant (2015 - 2016)***
    - Fast online estimation of RT60
    - Determine 'dry'/'wet' energy with short time window and hop size
    - development of VST plugins for realtime application

- ***Engine Order Cancellation (2014-2015)***
    - Implementation of in-vehicle engine noise cancellation
    - Multiple-speaker multiple mic configuration
    - Control state machine to allow distrubing events for example window opening
    - Deploying in many commercial car products

  This is the major project of the year. Engine noise is complicated mixture of harmonics of the rotor based frequency or RPM. We use adaptive filters to track the transfer function from the noise source to passengers, the estimated paths will be used to generate compensatory sounds from the speakers. The overall effect to every passenger is quiet zones surrounding individuals. We investigated the Matlab codes and made C/C++ implemnentations for DSP and simulation platform for tuning purpose. Supporting tooling features, for example, measuring the impulse responses of sencondary paths efficiently and accurately while without loss of user friendly, are also designed and implemented.

- ***Novel Control of Tracking Power Supply (2014 - 2015)***
    - creative method for power control of H-class PSU
    - save the cost of expensive DAC
    - used in amplifier products

- ***ACOspector(TM) – non-invasive fluid measurement in real-time (2010 - 2014)***

  Lead the development of the industrial fluid property measurement sensor based on active acoustic spectroscopy measurement. Products have been deployed among pulp and paper, chemical factories in Sweden. Some information can be found in www.acosense.com. Contributions:

  - made the first prototype that delivers to customers
  - improved the product to its second generation

- ***Matrix Power Converter Design and Implementation (2007 - 2008)***

  Matrix converter can be used as motor drivers(PSM, Induction Motor) thanks to its high power density. This is an academic project for my first graduate thesis. Matrix power converter has the merit of higher energy density compared to traditional back-to-back rectifierinverter topology. Nevertheless the synchronized control of all semiconductor switchers poses harder control problem against the old. In this National Science Foundation project, theory of matrix converters has been studied and a prototype has been engineered. The novel prototype is based on four-leg structure thus allows for more intuitive implementation in a carrier-modulation style.

  This is a national science foundation project thus the deliverables are publications below:

  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

Language Skills

1. Chinese - native
2. English - proficient
3. Swedish - basic

Driver's License - C1