

Srikanth Ronanki

Areas of interest

- Speech Synthesis/Recognition, Multimodal learning

Education

- Oct 2014 – **PhD, Centre for Speech Technology Research – University of Edinburgh.**
present Expected: June 2018
Adviser: Prof. Simon King
- July 2011 – **MS by Research, Speech and Vision Lab – IIIT-Hyderabad.**
Apr 2014 Adviser: Dr. Kishore S. Prahallad
Cumulative GPA: 8.25/10
- July 2007 – **Bachelor of Technology – IIIT-Hyderabad.**
July 2011 Cumulative GPA: 7.1/10

Technical Skill Set

- Languages Python, C, C++ (Basic), Java (Basic)
ML Tools Theano, Keras, PyTorch
Speech Tools Merlin, Festival, Festvox, Flite, HTK/HTS, Sphinx, Kaldi
Systems Linux, macOS, Microsoft Windows

Work experience

- Feb 2018 – **Amazon Internship (Applied Scientist), CAMBRIDGE, UK.**
present - Prototyping the next generation algorithms for Text-To-Speech, Alexa language technologies.
- Oct 2014 – **Centre for Speech Technology Research, EDINBURGH UNIVERSITY, UK.**
present - Co-author of CSTR's Merlin speech synthesis toolkit
- Participated in Blizzard challenge workshop for speech synthesis (2015-17)
- Implemented a hierarchical encoder-decoder model for statistical parametric speech synthesis
- Performed joint modeling of duration and acoustic parameters using LSTMs
- Merlin: <https://github.com/CSTR-Edinburgh/merlin>
- Github: <https://github.com/ronanki>
- Feb–May 2017 **NII Internship, TOKYO, Japan.**
- Investigated an approach for hybrid speech synthesis based on seq2seq modeling.
- Proposal: http://srikanthr.in/Hybrid_speech_synthesis_presentation.pdf
- Jan–Oct 2014 **[24]7 inc., ILABS, Bangalore, India.**
- Implemented text normalization for chat data and integrated with virtual assistant for chats.
- Investigated the importance of use of Omnichannel data for natural language understanding and explored different machine learning techniques for modeling.
- Experimented with fine-tuning of parameters in statistical language models for speech recognition.
- Framework: *Python/Java environment*, Tools: *SRISLMTK*;

- May–Sep 2013 **Google Summer of code Internship**, ANKUR-INDIA, India.
- Developed speech based Indic IVRS with open-source Ankur-India organization.
 - Constructed wrappers which work around Festival/Festvox for TTS and Sphinx/CMUCLMTK for ASR.
 - Implemented a small module for speech enhancement and tested the whole system on limited vocabulary domain-specific data.
 - Project URL: <http://indicivrs.blogspot.com>
 - Code: https://github.com/bhavibond/Indic_IVRS
- Feb–May 2013 **CSTR Internship**, EDINBURGH UNIVERSITY, UK.
- Implemented an approach for clustering of syllables with prosodic information and integrated with HMM models for acoustic modeling.
- July–September 2012 **Google Summer of code Internship**, CMUSPHINX, USA.
- Designed a web based Pronunciation Evaluation scoring routine using CMUSphinx speech recognition which provides necessary feedback on mispronunciation at phone/word level.
 - Implemented text-independent scoring routine and evaluated on Indian-accented speech data.
 - Project URL: <http://pronunciationeval.blogspot.com>
 - Code: <http://goo.gl/YUAEPg>

Publications

1. **Srikanth Ronanki**, Oliver Watts, Simon King. “A Hierarchical Encoder-Decoder Model for Statistical Parametric Speech Synthesis”. In proc. of Interspeech, 2017. [pdf]
2. **Srikanth Ronanki**, Manuel Sam Ribeiro, Felipe Espic, Oliver Watts. “The CSTR entry to the Blizzard Challenge 2017”. In proc. of Blizzard Challenge Workshop, 2017. [pdf]
3. **Srikanth Ronanki**, Oliver Watts, Simon King, Gustav Eje Henter. “Median-Based Generation of Synthetic Speech Durations using a Non-Parametric Approach”. In proc. of IEEE-SLT, 2016. [pdf]
4. Thomas Merritt, **Srikanth Ronanki**, Zhizheng Wu, Oliver Watts. “The CSTR entry to the Blizzard Challenge 2016”. In proc. of Blizzard Challenge Workshop, 2016. [pdf]
5. **Srikanth Ronanki**, Zhizheng Wu, Oliver Watts, Simon King. “A Demonstration of the Merlin Open Source Neural Network Speech Synthesis System”. In proc. of SSW9, 2016. [pdf]
6. **Srikanth Ronanki**, Siva Reddy, Bajibabu Bollepalli, Simon King. “DNN-based speech synthesis for Indian languages from ASCII text”. In proc. of SSW9, 2016. [pdf]
7. **Srikanth Ronanki**, Gustav Eje Henter, Zhizheng Wu, Simon King. “A template-based approach for speech synthesis intonation generation using LSTMs”. In proc. of Interspeech, 2016. [pdf]
8. Gustav Eje Henter, **Srikanth Ronanki**, Oliver Watts, Mirjam Wester, Zhizheng Wu, Simon King. “Robust TTS duration modelling using DNNs”. In proc. of ICASSP, 2016. [pdf]
9. Oliver Watts, **Srikanth Ronanki**, Zhizheng Wu, Tuomo Raitio, Antti Suni. “The NST–GlottHMM entry to the Blizzard Challenge 2015”. The Blizzard Challenge workshop, 2015. [pdf]
10. **Srikanth Ronanki**, Li-Bo, James Salsman. “Automatic Pronunciation Evaluation And Mispronunciation Detection Using CMUSphinx”. In proc. of SLP-TED workshop, Coling-2012. [pdf]
11. **Srikanth Ronanki**, Bajibabu Bollepalli and Kishore S. Prahallad. “Duration Modelling in Voice Conversion Using Artificial Neural Networks”. In proc. of IWSSIP, 2012. [pdf]
12. Bajibabu Bollepalli, **Srikanth Ronanki**, Sathya Adithya Thati, Bhiksha Raj, B. Yegnanarayana and Kishore S. Prahallad. “A Comparison of Prosody Modification Using Instants of Significant Excitation and Mel-cepstral Vocoder”. In proc. of Centenary Conference, IISc Bangalore, 2011. [pdf]

Other technical papers

1. **Srikanth Ronanki**, Gustav Eje Henter, Zhizheng Wu, and Simon King, “*A template-based approach for intonation generation using LSTMs*”, UKSpeech workshop, 2016.
2. **Srikanth Ronanki**, Zhizheng Wu, and Robert A. J. Clark, “*Joint Modeling of F0 and Duration in Deep Neural Network Based Speech Synthesis*”, UKSpeech workshop, 2015.
3. **Srikanth Ronanki**, Oliver Watts, Simon King and Robert A. J. Clark, “*Syllable based models for prosody modeling in HMM based speech synthesis*”, Simple4All intern report, 2013.
4. **Srikanth Ronanki**, Kishore S. Prahallad, “*Prosody Modeling for Voice Conversion*”, research project report, 2013.

Conferences/Workshops attended

Conferences Interspeech 2016-17, SICSA PhD 2016, Coling 2012, PAKDD 2010

Workshops SSW9 2016, Blizzard Challenge 2015-17, UKSpeech 2015-17

Thesis work

- Masters Primarily, worked towards speech synthesis systems to overcome some of the existing drawbacks.
- Thesis Explored the issues in modeling the sub-word units, and proposed an approach using longer size units such as syllable, and build a statistical template for each syllable using dynamic programming to inherently capture the trajectories. [pdf]

Extra-Curricular activities

- Students Finance Secretary of IIIT-H from 2009-11
- System Administrator for Electronic Courier Portal in IIIT-H from 2009-2012
- Member of a Volleyball Team and Football at Inter-college level

References

Available upon request